

# **Identification of *Tb*RP2-interacting proteins using proximity-dependent biotinylation (BioID)**

**Xin Qi**

**This thesis is submitted in partial fulfilment of the requirements for the degree of  
Doctor of Philosophy to**

**Lancaster University**

**July 2017**

# Abstract

The protein RP2 is a tubulin cofactor C-domain containing protein with important roles in ciliogenesis. In humans, mutations in the RP2 gene are associated with 10-15% of cases of X-linked retinitis pigmentosa; a devastating disease characterised by progressive degeneration of retinal photoreceptors. Although XRP2 was initially proposed to function as a GTPase-activating protein (GAP) for tubulin, evidence now suggests that it acts as a GAP for Arl3 (a small GTPase) and together with Arl3 is involved in trafficking proteins to the cilium. I have been studying RP2 function in the flagellated protist *Trypanosoma brucei*, a tractable model to study eukaryotic flagellum assembly but also a parasite of medical and veterinary importance in sub-Saharan Africa. Thus, the study of RP2 in trypanosomes has relevance for parasitology, but also the human inherited disease retinitis pigmentosa. However, important differences exist between XRP2 and *TbRP2*, for instance *TbRP2* lacks the consensus sequence specifying N terminal myristoylation (a modification that targets XRP2 to the basal body in mammalian cells) but rather utilises twinned TOF-LisH motifs at the N-terminus of the protein to direct basal body targeting. To further interrogate the targeting and function of *TbRP2*, I employed proximity-dependent biotin identification (BioID), in combination with quantitative proteomic (SILAC) techniques, to identify putative *TbRP2*-interacting proteins *in vivo*. A selected cohort of these proteins were subsequently interrogated by bioinformatics, localised within the cell using a PCR only (pPOT) YFP-tagging strategy and their potential roles in flagellum formation interrogated using inducible RNA interference (RNAi). My studies identified: (i) an Arl3-related protein as the likely molecular client of *TbRP2* GAP activity; (ii) the trypanosome mature basal body as a hub for molecular chaperone activity associated with eukaryotic flagellum assembly; and (iii) insight into lineage-specific aspects of basal body biogenesis, as illustrated by the unusual spatial and temporal inheritance of large, trypanosomatid-specific protein of unknown function (*TbBBP590*).

# Acknowledgements

I would like to say thank you to my supervisors, Paul McKean and Michael Ginger for all their support, guidance, ideas and opportunities they have provided throughout my time at Lancaster. I am so grateful to be a member of their lab, they are the best supervisors I could ever asked for. I would also like to say a special thanks to Dr. Jane Harmer for all her help in the lab from the beginning until the end, and her patience and generosity.

Thank you to Dr Mick Urbaniak, Dr Corinna Benz, Dr Derek Gatherer, Dr Fiona Benson and Janice Drinkall for their help and advice, and their kind generosity with reagents and equipment.

A huge thank you to my lovely friends from the lab over the years! Kurimun, Jon, Louise, Camila, Harriet, Alex, Ping, Urvi. You guys are awesome and are responsible for some of the best times at Lancaster. And a special thanks to my wonderful friends Amber Lynch and Grupreet Patel got me through the harder times, so lucky to have you guys in my life.

This project was partially funded by the Faculty of Health and Medicine, and this financial support was invaluable to me. However, my living cost was provided by my spectacular parents, Ping Qi and Sheng-min Zheng. I would not be in the UK and where I am today without their help, support. Thank you so much and I love you both. Finally, I would love to say thank you to my fiancé Lorenzo Zaccagnini for his enormous amounts of support and encouragement, and for being so caring, patient with me throughout these challenging years!

## **Declaration**

I declare that this thesis was composed by myself and has not been submitted in substantially the same form for the award of a higher degree elsewhere.



# Table of Contents

<b>Abstract .....</b>	<b>i</b>
<b>Acknowledgements .....</b>	<b>ii</b>
<b>Declaration .....</b>	<b>iii</b>
<b>Table of Contents .....</b>	<b>iv</b>
<b>List of Figures .....</b>	<b>x</b>
<b>List of Tables .....</b>	<b>xvi</b>
<b>Abbreviations.....</b>	<b>xviii</b>
<b>Chapter 1 Introduction .....</b>	<b>1</b>
1.1 Cilium/Flagellum.....	1
1.1.1 General background .....	1
1.1.2 Cilium/Flagellum structure .....	3
1.1.3 Cilium biogenesis .....	6
1.1.4 Ciliary signalling .....	7
1.1.5 IFT machinery.....	8
1.2 The base of the cilium .....	15
1.2.1 The ciliary gate .....	15
1.2.2 The ciliary gate and ciliopathies .....	17
1.2.2.1 NPHP 1-4-8 protein complex.....	22
1.2.2.2 MKS/B9 protein complex.....	24
1.2.2.3 NPHP 5-6/CEP290 protein complex.....	27
1.2.2.4 Inversin compartment.....	29
1.2.2.5 Other proteins and lipids.....	30
1.3 <i>Trypanosoma brucei</i> – a model system for studying flagellum biogenesis .....	32
1.3.1 Background and life cycle .....	32
1.3.2 Cell Architecture .....	34

1.3.3	Flagellum biogenesis and cell cycle .....	36
1.3.4	The Importance of the <i>T. brucei</i> flagellum .....	42
1.3.4.1	Cell morphogenesis and cell division .....	42
1.4	Tubulin folding pathway .....	44
1.5	TBCC domain containing proteins .....	46
1.5.1	Canonical TBCC .....	46
1.5.2	TBCCD1 .....	47
1.5.3	XRP2 .....	49
1.5.4	TbRP2 .....	55
1.6	Project aims .....	57
<b>Chapter 2</b>	<b>Material and Methods .....</b>	<b>58</b>
2.1	Chemicals and Reagents .....	58
2.2	Solutions and Buffers .....	60
2.2.1	Solutions and Buffers for Mass Spectrometry (BioID/SILAC) (OB-on beads digest, IG-in gel digest) .....	62
2.3	Antibiotics and drug stock solutions .....	63
2.4	Vector maps .....	64
2.5	Antibodies and PCR primers .....	67
2.6	Oligonucleotide primers .....	68
2.7	PCR amplification .....	74
2.8	Purification of PCR products .....	75
2.9	DNA Ligations .....	75
2.10	Transformation into Competent <i>E. coli</i> XL1-Blue cells .....	75
2.11	Plasmid purification using a Miniprep kit .....	76
2.12	Gel extraction and purification of DNA .....	76
2.13	Restriction enzyme digests .....	76

2.14 Agarose gel electrophoresis .....	76
2.15 Preparation of plasmid DNA for transfection into <i>T. brucei</i> .....	77
2.16 <i>T. brucei</i> cell culture .....	77
2.17 Transfection of DNA into <i>T. brucei</i> .....	77
2.18 Glycerol stock for <i>T. brucei</i> cultures .....	78
2.19 Sample preparation for mass spectrometry (BioID-SILAC) .....	78
2.19.1 Culturing of <i>T. brucei</i> in SILAC media .....	78
2.19.2 BioID Sample Preparation (on beads digest) .....	79
2.19.3 BioID Sample Preparation (in gel digest) .....	81
2.20 Liquid chromatography mass spectrometry .....	83
2.21 Analysis of mass spectrometry data .....	83
2.22 Preparation of slides for fluorescence microscopy .....	84
2.23 Probing slides for fluorescence microscopy .....	84
2.24 Fluorescence microscopy .....	85
2.25 Preparation of <i>T. brucei</i> protein samples .....	85
2.26 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE) .....	85
2.27 Immunoblotting .....	85
<b>Chapter 3 Identification of putative <i>TbRP2</i> interacting/near neighbour proteins using proximity-dependent biotinylation .....</b>	<b>87</b>
3.1 Introduction .....	87
3.2 BioID in theory .....	88
3.3 Generation of plasmids to allow expression of <i>TbRP2::myc::BirA*</i> fusion proteins .....	90
3.4 Localisation of <i>TbRP2::myc::BirA*</i> expressing protein in <i>T. brucei</i> .....	92
3.5 Analysis of expressing <i>TbRP2::myc::BirA*</i> and <i>TbRP2<sup>Δ134-463</sup>::myc::BirA*</i> in <i>T. brucei</i> by immunoblotting .....	102

3.6 Stable Isotope Labelling with Amino acid in the Cell culture (SILAC) .....	104
3.7 Validation of <i>Tb</i> RP2-interacting proteins and near neighbours by BioID and SILAC .....	106
3.7.1 Experiment 1 - Whole cell -1, in gel digestion .....	108
3.7.2 Experiment 2 - Cytoskeletons, in solution/on bead digestion .....	111
3.7.3 Experiment 3- Whole cell - 2, in solution/on bead digestion .....	118
3.8 Summary .....	127
<b>Chapter 4 Bioinformatics validation .....</b>	<b>128</b>
4.1 Introduction.....	128
4.2 Heat shock protein 70 .....	132
4.3 ADP-ribosylation factor-like protein 3-Like (ARL3L) protein.....	141
4.4 Kinetoplastid membrane protein-11 (KMP-11) .....	145
4.5 Microtubule-associated repetitive protein (MARF).....	146
4.6 <i>Tb</i> BBP590 .....	147
4.7 Cytoskeletal-Associated Protein (CAP) 51 .....	151
4.8 VHS (VPS-27, Hrs and STAM)-domain containing protein .....	154
4.9 p25- $\alpha$ /TPPP (Tubulin-Polymerization-Promoting Protein) .....	158
4.10 TCP-1- $\epsilon$ (T-Complex Protein 1 subunit epsilon) .....	163
4.11 Protein Kinase ( <i>Tb</i> 927.9.6560) .....	166
4.12 ARL2-binding protein (BART, Binder of Arl Two) like 1 .....	176
<b>Chapter 5 Localisation of putative <i>Tb</i>RP2 interacting /near neighbour proteins..</b>	<b>180</b>
5.1 Introduction.....	180
5.2 Generation of procyclic cell lines expressing YFP-tagged proteins.....	180
5.3 ADP-ribosylation factor-like protein 3 Like (ARL3 L) .....	183
5.4 Kinetoplastid membrane protein 11 (KMP-11).....	185
5.4.1 Further interrogation of <i>Tb</i> KMP-11::YFP flagellum localisation .....	185

5.5 Binder of Arl2 (BART) Like 1 .....	190
5.6 Tubulin Polymerisation Promoting Protein (TPPP)/p25- $\alpha$ .....	192
5.7 T-complex protein 1 epsilon (TCP-1- $\epsilon$ ) .....	194
5.8 VHS-domain containing protein .....	196
5.9 Heat Shock Protein 70 (HSP70) .....	199
5.10 Protein Kinase (PK) ( <i>Tb</i> 927.9.6560) .....	202
5.11 Microtubule associated repetitive protein (MARF) ( <i>Tb</i> 927.10.10360) .....	204
5.12 <i>Tb</i> BBP590 .....	208
5.13 <i>Tb</i> Cytoskeletal-associated protein (CAP51) ( <i>Tb</i> 927.7.2640) .....	213
5.14 Summary .....	215
<b>Chapter 6 Further characterisation of <i>Tb</i>RP2-proximal proteins .....</b>	<b>217</b>
6.1 Introduction.....	217
6.2 <i>Tb</i> KMP-11 .....	217
6.2.1 <i>Tb</i> KMP-11 flagellum localisation – relationship to FLAM3.....	221
6.2.2 <i>Tb</i> KMP-11 flagellum localisation – relationship to calmodulin .....	224
6.3 <i>Tb</i> Cytoskeletal-associated protein (CAP51) .....	227
6.4 T-complex protein 1 subunit epsilon (TCP-1- $\epsilon$ ).....	232
6.5 ADP-ribosylation factor-like 3 (ARL3)-like .....	238
6.6 ARL2-binding protein (BART, Binder of ARL Two) like-1 .....	242
6.7 Heat Shock Protein 70 (HSP70) .....	246
6.8 Vps-27, Hrs and STAM (VHS) domain containing protein .....	250
6.10 Microtubule-associated repetitive protein (MARF).....	258
6.11 <i>Tb</i> BB590 .....	260
6.12 Tubulin polymerization promoting protein (TPPP/P25- $\alpha$ ).....	262
6.13 Effect of <i>Tb</i> RP2 RNAi on the expression and localisation of selected putative <i>Tb</i> RP2-interacting/near neighbour proteins.....	266

6.13.1 KMP-11.....	266
6.13.2 HSP70.....	269
6.13.3 <i>Tb</i> BBP590.....	272
6.13.4 BART LIKE 1 .....	275
6.13.5 ARL3 Like.....	279
6.14 Summary .....	283
<b>Chapter 7 Discussion .....</b>	<b>286</b>
<b>Chapter 8 Future Work .....</b>	<b>305</b>
<b>Chapter 9 Appendices.....</b>	<b>310</b>
<b>Appendix I.....</b>	<b>310</b>
<b>Appendix II.....</b>	<b>314</b>
<b>Appendix III.....</b>	<b>335</b>
<b>References .....</b>	<b>349</b>

# List of Figures

Figure 1.1 A diversity of cilia and flagella can be found in eukaryotic organisms. ....	2
Figure 1.2 The general structure and the cross section of a cilium.....	5
Figure 1.3 Ciliogenesis; how cilium/flagellum assembled. ....	7
Figure 1.4 The assembly and trafficking of the primary cilium by intraflagellar transport (IFT). ....	10
Figure 1.5 Model organization chart of four protein complexes and the possible connections.....	19
Figure 1.6 The life cycles of <i>T.brucei</i> in tsetse fly and human host .....	33
Figure 1.7 Cartoon showing the cellular morphology of a procyclic form <i>T. brucei</i> cell .....	35
Figure 1.8 The trypanosome flagellum in procyclic form. ....	40
Figure 1.9 Cartoon showing the timing and order of the major morphological events during the cell cycle in procyclic <i>T. brucei</i> . ....	41
Figure 1.10 Schematic diagram shows the canonical tubulin-folding pathway. ....	45
Figure 1.11 Diagram shows the domain structure of RP2 protein. ....	49
Figure 2.1 Vector maps. ....	66
Figure 3.1 Schematic architectures and localization of C-terminally deleted <i>TbRP2</i> ..	88
Figure 3.2 Schematic showing the application of the BioID method .....	89
Figure 3.3 Schematic of DNA construct of <i>TbRP2</i> <sup>Δ134-463</sup> ::myc::BirA* in pEX-K4 vector. ....	90
Figure 3.4 Gel analysis of restriction digested <i>TbRP2</i> <sup>Δ134-463</sup> ::myc::BirA* in pDex377 clones. ....	91
Figure 3.5 The diagram shows the purification procedures to determine the localisation and expressions of <i>TbRP2</i> ::myc::BirA*. ....	92
Figure 3.6 Immunofluorescence analysis of a detergent extracted cells expressing <i>TbRP2</i> <sup>Δ134-463</sup> ::myc::BirA* in the presence of excess biotin in detergent extracted fashion.....	94
Figure 3.7 Immunofluorescence analysis of detergent extracted cells expressing <i>TbRP2</i> <sup>Δ134-463</sup> ::myc::BirA* probed with TRITC-conjugated streptavidin .....	95

Figure 3.8 Immunofluorescence analysis of a detergent extracted cells expressing <i>TbRP2::myc::BirA*</i> in the presence of excess biotin in detergent extracted fashion..	96
Figure 3.9 Immunofluorescence analysis of detergent extracted cells expressing <i>TbRP2::myc::BirA*</i> probed with TRITC-conjugated streptavidin .....	97
Figure 3.10 Immunofluorescence analysis of a cell inducible expressing <i>TbRP2<sup>Δ32-463</sup>::myc::BirA*</i> in the presence of excess biotin in detergent extracted fashion.....	98
Figure 3.11 Immunofluorescence analysis of detergent extracted cells expressing <i>TbRP2<sup>Δ322-463</sup>::myc::BirA*</i> probed with TRITC-conjugated streptavidin .....	99
Figure 3.12 Immunofluorescence analysis of a detergent extracted cells expressing <i>TbRP2<sup>Δ171-463</sup>::myc::BirA*</i> in the presence of excess biotin in detergent extracted fashion .....	100
Figure 3.13. Immunofluorescence analysis of detergent extracted cells expressing <i>TbRP2<sup>Δ171-463</sup>::myc::BirA*</i> probed with TRITC-conjugated streptavidin .....	101
Figure 3.14 Expression Immunoblots of <i>TbRP2<sup>Δ134-463</sup>::myc-BirA*</i> and <i>TbRP2 FL::myc::BirA*</i> .....	103
Figure 3.15 SILAC workflow. ....	105
Figure 3.16 Example of a LC-MS/MS scatter plot showing the proteins identified in the SILAC-enriched ‘medium or heavy’ cell line .....	107
Figure 3.17 LC-MS/MS scatter plot comparing proteins identified in the SILAC-enriched in <i>TbRP2<sup>Δ134-463</sup>/PTPs</i> vs <i>TbRP2/PTPs</i> in whole cells .....	109
Figure 3.18 Scatter plot showing proteins identified in the SILAC-enriched <i>TbRP2<sup>Δ134-463</sup>/PTPs</i> or <i>TbRP2/PTPs</i> vs intensity in cytoskeleton cells .....	115
Figure 3.19 LC-MS/MS scatter plot comparing proteins identified in the SILAC-enriched <i>TbRP2<sup>Δ134-463</sup>/PTPs</i> vs <i>TbRP2/PTPs</i> .....	116
Figure 3.20 The combined SILAC ratio of each selected proteins in Cyto expt.....	117
Figure 3.21 Scatter plot showing proteins identified in the SILAC-enriched <i>TbRP2 (M)/PTPs</i> vs <i>TbRP2 (H)/PTPs</i> in whole cells .....	123
Figure 3.22 The LC-MS/MS scatter plot comparing proteins identified in the SILAC-enriched <i>TbRP2 (M)/PTPs</i> or <i>TbRP2 (H)/PTPs</i> vs intensity .....	124
Figure 3.23 The LC-MS/MS scatter plot comparing proteins identified in the SILAC-enriched in Cyto expt and WC expt .....	125



Figure 3.24 The SILAC ratio of each selected proteins in the BioID .....	126
Figure 4.1 Schematic cartoon HSP70 domain structures and multiple sequence alignments between human and <i>T. brucei</i> .....	135
Figure 4.2 Schematic cartoon of <i>T. brucei</i> HSP70 domain structures and multiple sequence alignments .....	140
Figure 4.3 Schematic cartoon of ARL3 domain structures and multiple sequence alignments between human and <i>T. brucei</i> .....	143
Figure 4.4 Schematic cartoon of <i>T. brucei</i> ARL3 domain structures and multiple sequence alignments .....	144
Figure 4.5 Schematic cartoon of Kinetoplastids KMP-11 domain structures and multiple sequence alignments .....	145
Figure 4.7 Interpro Scan and IUPRED analysis of <i>TbBBP590</i> .....	150
Figure 4.8 Interpro Scan and IUPRED analysis of <i>TbCAP51</i> .....	153
Figure 4.9 Schematic cartoons of VHS domain structure .....	157
Figure 4.10 Schematic cartoons of p25- $\alpha$ domain structure.....	162
Figure 4.11 Schematic cartoons shows the domain structures of TCP-1- $\epsilon$ .....	165
Figure 4.12 Schematic cartoons shows the domain structure of protein kinase .....	175
Figure 4.13 Schematic cartoons shows the domain structure of BARTL1 proteins ..	179
Figure 5.1 PCR products of all selected candidate proteins .....	181
Figure 5.2 Confirmation of the expression of YFP-tagged proteins by immunoblotting. ....	182
Figure 5.3 Immunofluorescence images showing the localisation of <i>TbARL3L::YFP</i> in procyclic form <i>T. brucei</i> .....	184
Figure 5.4 Fluorescence images showing localisation of <i>TbKMP-11::YFP</i> in cells at different stages of the cell cycle. ....	187
Fig. 5.5 Immunofluorescence images of cells expressing <i>TbKMP-11::YFP</i> co-labelled with the anti-PFR antibody L8C4. ....	188
Fig. 5.6 Immunofluorescence images of <i>T. brucei</i> procyclic cells showing localisation of <i>TbKMP-11::YFP</i> in relation to the flagellar connector antibody AB1 .....	189
Figure 5.7 Immunofluorescence images showing the localisation of <i>TbBARTL1::YFP</i> in procyclic form <i>T. brucei</i> .....	191

Figure 5.8 Immunofluorescence images showing the localisation of <i>TbP25-α::YFP</i> in procyclic form <i>T. brucei</i> .....	193
Figure 5.9 Immunofluorescence images showing the localisation of <i>TbTCP-1-ε::YFP</i> in procyclic form <i>T. brucei</i> .....	195
Figure 5.10 Immunofluorescence images showing the localisation of <i>TbVHS::YFP</i> in procyclic form <i>T. brucei</i> .....	197
Figure 5.12 Immunofluorescence images showing the localisation of <i>TbHSP70::YFP</i> in procyclic form <i>T. brucei</i> .....	200
Figure 5.13 Immunofluorescence images showing the localisation of <i>TbHSP70::YFP</i> in procyclic form <i>T. brucei</i> .....	201
Figure 5.14 Immunofluorescence images showing the localisation of <i>TbPK::YFP</i> in procyclic form <i>T. brucei</i> .....	203
Figure 5.15 Immunofluorescence images showing the localisation of <i>TbMARP::YFP</i> in procyclic form <i>T. brucei</i> (WC).....	205
Figure 5.16 Immunofluorescence images showing the localisation of <i>TbMARP::YFP</i> in procyclic form <i>T. brucei</i> (Cyto) .....	206
Figure 5.17 Immunofluorescence images showing the localisation of <i>TbMARP::YFP</i> in procyclic form <i>T. brucei</i> (Isolated Flagella) .....	207
Figure 5.18 Immunofluorescence images showing the localisation of <i>TbBBP590::YFP</i> in procyclic form <i>T. brucei</i> (Cyto). .....	210
Figure 5.19 Immunofluorescence images showing the localisation of <i>TbBBP590::YFP</i> in procyclic form <i>T. brucei</i> (Isolated Flagella). .....	212
Figure 5.20 Immunofluorescence images showing the localisation of <i>TbCAP51::YFP</i> in procyclic form <i>T. brucei</i> .....	214
Figure 5.21 Schematic summarising localisations of putative <i>TbRP2</i> -interacting proteins and nearby neighbours.....	215
Figure 6.1 RNA interference-mediated ablation of <i>TbKMP-11</i> results in multiple defects in procyclic form cells.....	220
Figure. 6.2 The expression and localisation of <i>TbKMP-11</i> is not dependent upon FLAM 3 .....	223
Figure 6.3 The level of <i>TbKMP-11::YFP</i> expression was examined following CaM RNAi induction .....	226

Figure.6.5 RNA interference-mediated ablation of <i>TbCAP51</i> results in multiple defects in procyclic form cells.....	231
Figure 6.6 RNA interference-mediated ablation of <i>TbTCP-1-ε</i> results in multiple defects in procyclic form cells.....	237
Figure 6.7 The effect of RNA interference-mediated ablation of <i>TbARL3L</i> on procyclic form cells.....	241
Figure 6.8. The effect of RNA interference-mediated ablation of <i>TbBARTL1</i> on procyclic form cells. ....	245
Figure 6.9. The effect of RNA interference-mediated ablation of <i>TbHSP70</i> on procyclic form cells. ....	249
Figure 6.10 The effect of RNA interference-mediated ablation of <i>TbVHS</i> on procyclic form cells.....	253
Figure 6.11 The effect of RNA interference-mediated ablation of <i>TbPK</i> on procyclic form cells.....	257
Figure 6.13 The effect of RNA interference-mediated ablation of <i>TbBBP590</i> on procyclic form cells. ....	261
Figure 6.14 The effect of RNA interference-mediated ablation of <i>TbP25-α</i> on procyclic form cells. ....	265
Figure 6.15 The effect of <i>TbRP2</i> RNAi mediated ablation on <i>TbKMP-11</i> localization and expression. ....	268
Figure 6.16 The effect of <i>TbRP2</i> RNAi mediated ablation on <i>TbHSP70</i> localisation and expression. ....	271
Figure 6.17 The effect of <i>TbRP2</i> RNAi mediated ablation on <i>TbBBP590</i> localisation and expression. ....	274
Figure 6.18 The effect of <i>TbRP2</i> RNAi mediated ablation on <i>TbBARTL1</i> localisation and expression. ....	278
Figure 6.19 The effect of <i>TbRP2</i> RNAi mediated ablation on <i>TbARL3L</i> localisation and expression. ....	282
Figure 7.1. Working model of the role of Arls and Rabs in the cilium.....	292
Figure 7.2 Schematic diagram of the recruitment of <i>TbBBP590</i> in <i>T. brucei</i> procyclic (insect) form cells during cell division cycle.....	299
Figure 7.3 Potential interactions and localisations of selected FAZ proteins.....	302

Figure 7.4 Working models of possible <i>Tb</i> RP2 interacting/nearby partners .....	304
Figure 8.1 Proposed interaction/localisation model of <i>ruvBL</i> and <i>pih1/Nop17</i> proteins to <i>Tb</i> RP2 and in HSP70/90-mediated axonemal dynein assembly, based on the studies in other models. ....	306

# List of Tables

Table 1.1 Components of the IFT proteins in different organisms.....	14
Table 1.2 List of overview of human ciliopathic-associated genes encode proteins ..	20
Table 2.1 Recipe for making SDM-79-SILAC media for BioID experiments .....	63
Table 2.2 Recipe for making light-, medium- and heavy-labelled amino acids.....	63
Table 2.3 Antibiotics and drugs used during this study .....	64
Table 2.5 Antibodies used for immunofluorescence microscopy .....	67
Table 2.6 Antibodies used for Immunoblotting detection .....	67
Table 2.7 Streptavidin conjugated protein used for IF and IB .....	68
Table 2.8 pPOT primer sequences used in PCR reactions .....	69
Table 2.9 Sequences of oligonucleotide primers used to generate constructs for RNAi knockdown experiments.....	72
Table 2.10 Standard PCR program. ....	74
Table 2.11 High fidelity PCR program. ....	74
Table 3.1 Summary of SILAC experiments performed in this thesis.....	108
Table 3.2 Proteins identified in <i>TbRP2</i> BirA* fusion protein expressing <i>T. brucei</i> by BioID-SILAC (WC-1).....	110
Table 3.3 Proteins identified in <i>TbRP2</i> BirA* fusion protein expressing <i>T. brucei</i> by BioID-SILAC (Cyto1+2) .....	114
Table 3.4 Proteins identified in <i>TbRP2</i> BirA* fusion protein expressing <i>T. brucei</i> by BioID-SILAC (WC-2) .....	120
Table 4.1 A broad range of microbial eukaryotes, including flagellated and non-flagellated organisms used in the protein analysis.....	129
Table 4.2 NCBI BLAST search of different organisms to the <i>T.brucei</i> BioID pull down proteins .....	130
Table 4.3 The <i>T. brucei</i> BioID pull down proteins compare with proteomic studies.....	130
Table 4.4 The HSP70 family of <i>Trypanosoma brucei</i> .....	139
Table 4.5 The orthologues of MARP in Kinetoplastids .....	146
Table 5.1 Table summarising the localisation(s) of selected putative <i>TbRP2</i> -interacting/near neighbour proteins. ....	216
Table 6.1 Localisation and phenotypic consequences of RNAi ablation .....	284

Table 6.2 Summary of experiments investigating if loss of <i>TbRP2</i> .....	285
Table 7.1 Summary table of all the proteins selected for further studies in <i>T. brucei</i> following BiolD analysis of <i>TbRP2::BirA</i> .....	288
Table 7.2 Some of the ciliary Arls, Rabs and their possible interactors. ....	293
Table 8.1 Selected proteins with good SILAC ratio (above 1), but their functionality did not investigate in this study. ....	308

# Abbreviations

a.a. Amino acids

ABC Ammonium bicarbonate

ARL ADP-ribosylation factor (ARF)-like

ARF ADP-ribosylation factor

APS Ammonium persulphate

BB Basal body

BBS Bardet–Biedl syndrome

BioID Proximity-dependent biotinylation identification

BLAST Basic local alignment tool

BSA Bovine serum albumin

BSF Bloodstream form

CAP Cytoskeletal associated protein

CC Coiled coil

CCT Cytosolic chaperonin

CYC Cyclin

DAPI 4', 6-diamidino-2-phenylindole

DAVs Distal appendage vesicles

DHC Dynein heavy chain

DLC Dynein light chain

DNA Deoxyribonucleic acid

DMSO Dimethyl sulfoxide

DTT Dithiothreitol

EB Elution buffer

EDTA Ethylene diamine tetra-acetic acid

EM Electron microscopy

ER Endoplasmic reticulum

FAZ Flagellum attachment zone

FC Flagella connector

FITC Fluorescein isothiocyanate

FP Flagellar pocket

FPC Flagellar pocket collar

GA Golgi apparatus

GAP GTPase-activating protein

GDP Guanosine diphosphate

GFP Green fluorescent protein

GTP Guanosine triphosphate



HAT Human African trypanosomiasis

Hh Hedgehog

HRP Horse radish peroxidase

IAA Iodoacetamide

IFT Intraflagellar transport

IG In-gel

JBTS Joubert Syndrome

KAP Kinesin-associated protein

kb Kilobase

kDa Kilo-dalton

kDNA kinetoplast DNA

KO Knockout (-/-)

LB Luria broth

LC-MS/MS Liquid chromatography tandem mass spectrometry

MAP Microtubule associated protein

MKS Meckel–Gruber syndrome

MT Microtubule

MTOC Microtubule organising centre

MtQ Microtubule quartet

MS Mass spectrometry

MW Molecular weight

NCBI National Centre for Biotechnology Information

NF New flagellum

NPHP Nephronophthisis

OF Old flagellum

PBS Phosphate buffered saline

PCF Procyclic form

PCR Polymerase chain reaction

PEME Pipes, EDTA, MgCL<sub>2</sub>, EGTA

PFR Paraflagella rod

PF Protofilaments

PKD *Polycystic kidney disease*

PTM Post translational modification

RP Retinitis pigmentosa

PPB Preprophase band

RNA Ribonucleic acid

RNAi RNA interference

SDS sodium dodecyl sulphate

SDS-PAGE sodium dodecyl sulphate polyacrylamide gel electrophoresis

SEM Scanning electron microscopy

SILAC Stable isotope labelling by amino acids in cell culture

Smo Smoothened

S/N Supernatant

SPMT *Subpellicular microtubule*

TAC Tripartite attachment complex

TAE Tris(hydroxymethyl)aminomethane-acetate ethylenediaminetetraacetic acid

TBCC Tubulin cofactor C

TFA Trifluoroacetic acid

TbFP *Trypanosoma brucei* flagella proteome

TBS Tris buffered saline

TCP Trypanosome cortex protein

TEM Transmission electron microscopy

TF Transitional fibre

TMEM Transmembrane

TRITC Tetramethylrhodamine isothiocyanate

Tris Tris(hydroxymethyl)aminomethane

TZ Transition zone

VSG Variant surface glycoprotein

YFP Yellow fluorescent protein

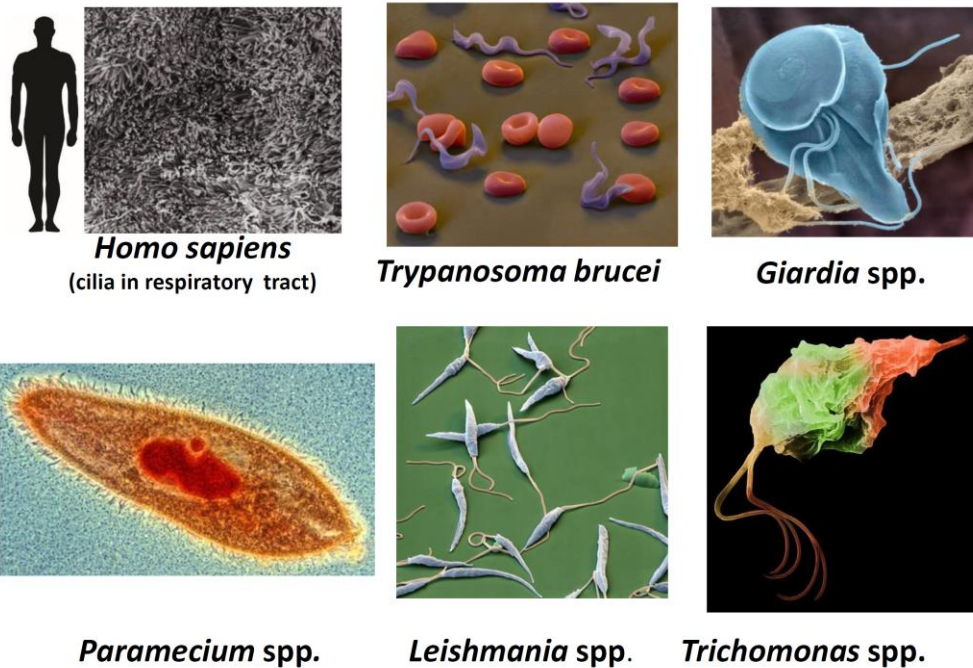
ZMG Zimmerman's postfusion medium

# Chapter 1 Introduction

## 1.1 Cilium/Flagellum

### 1.1.1 General background

Eukaryotic cilia/flagella (terms which will be used interchangeably in this chapter) are finger-like, microtubule-based structures, which protrude from the surface of many eukaryotic cells (Fig 1.1). The distinction between cilia and flagella is not always very clear. Historically, they have been distinguished by (1) their number - if there are many, (more than 100) they are normally called cilia, whereas if there are fewer than 8, they are usually called flagella, (2) their length - flagella are typically longer than cilia (Moran et al., 2014). Despite these differences, the basic organisation and mechanism of flagellum/cilium assembly are evolutionarily conserved and remarkably similar in phylogenetically diverse species. Both cilia/flagella have extremely important biological functions, ranging from motility to sensory perception and are vital to human health. Defects in cilium/flagellum functions can lead to a wide range of inherited human disorders called ciliopathies, which can affect respiratory system, retinal cells, kidney etc. (Tobin and Beales, 2009; Waters and Beales, 2011). Moreover, defects in cilia are also associated with other pathologies such as diabetes and cancer (Han and Alvarez-Buylla, 2010). In the following sections, I will provide details on cilium/flagellum structure and function as well as the mechanism(s) by which a cilium/flagellum is assembled.



**Figure 1.1 A diversity of cilia and flagella can be found in eukaryotic organisms.** Motile cilia in the **human respiratory tract** can move mucus and dirt out of human body. **Trypanosoma brucei** - the flagellum plays multiple roles (motility, host-parasite interaction, organelle positioning). **Giardia spp** - flagella are essential for motility, cell division and attachment to intestinal villi in the host body. **Paramecium** (a unicellular protozoa) is covered with many cilia whose main purpose is movement and the gathering of food particles. **Leishmania** is closely related to trypanosomes, and flagella can serve various roles, including sensory perception, cellular organisation and movement. **Trichomonas vaginalis** causes a sexually transmitted infection called Trichomoniasis, and flagella play an important role in motility. **Source: All images reprinted with permission from Encyclopaedia Britannica ImageQuest.**

### 1.1.2 Cilium/Flagellum structure

From a structural viewpoint, the cilium can be divided into sub-compartments that include the basal body, from where the microtubule axoneme is nucleated, the transition zone (TZ), and the core axoneme structure (Fig 1.2.A). At the basal body, there are nine radially organised microtubule triplets (designated the A, B and C tubules) (Fig 1.2.A-(1)). The A-tubule of each doublet is a complete MT, which consists of 13 protofilaments (pf), whereas the B and C-tubule is an incomplete MT often consisting of only 10 pfs (Nicastro et al., 2011; Sui and Downing, 2006). Because the cilium is a 'ribosome-free zone', in most flagellated eukaryotes proteins need to be delivered into the cilia by a process called intraflagellar transport (IFT); a motor driven process that transports proteins along axonemal microtubules. There are some species that do not require IFT for axoneme assembly, such as *Plasmodium falciparum*, which assemble the axoneme within the cytoplasm (Briggs et al., 2004a; Sinden and Smalley, 1976).

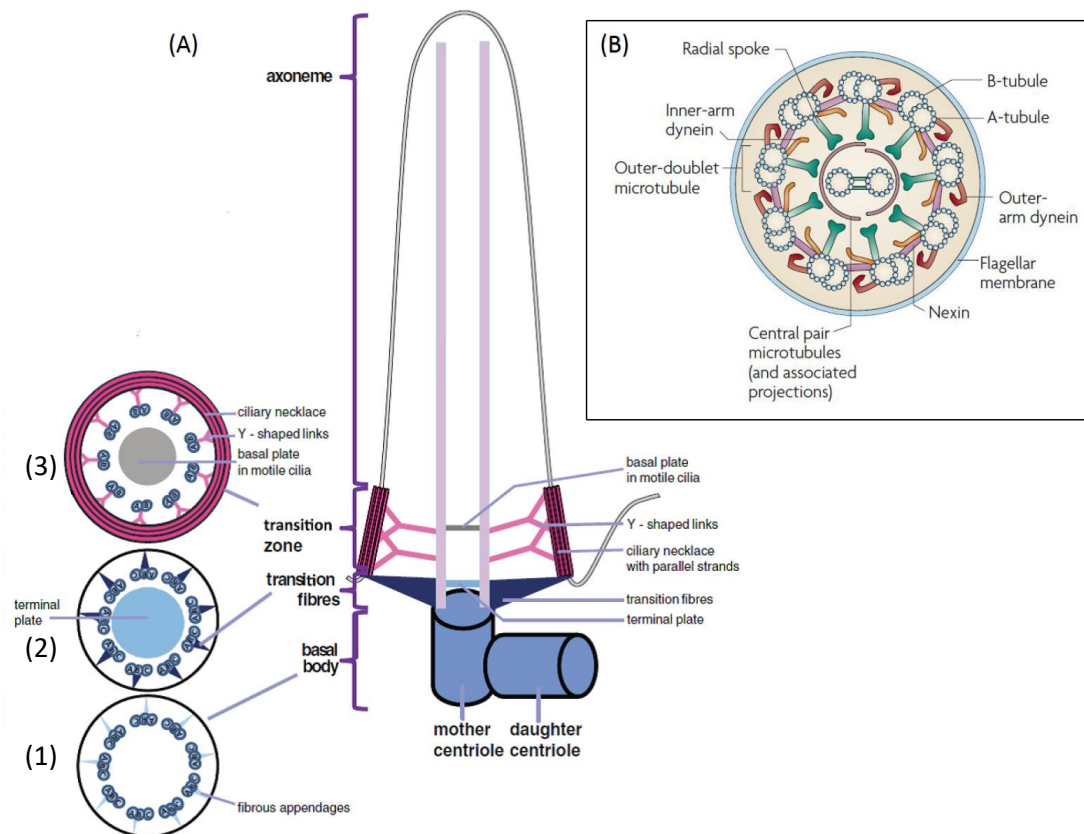
As the cilium extends to create a transition zone, the triplet microtubule structure of the basal body converts into doublet structure i.e. only the A and B tubules continue to extend to form the nine outer doublet microtubules of the axoneme shaft; the C tubules terminate within the transition zone between the basal body and the shaft. The proximal end of transitional fibres (Fig 1.2.A-(2)) connect the microtubule doublets to the membrane and mark the compartment border at which IFT proteins accumulate (Ishikawa and Marshall, 2011). The transition zone (TZ) is characterised by the Y-shaped linkers, which bridge from microtubule doublets to the ciliary membrane and distal appendages. Together the transitional fibres and the so-called ciliary necklace are proposed to form a physical and selective barrier (i.e. the ciliary gate) at the base of the cilium, which regulates the movement of ciliary proteins between the cytoplasm and ciliary compartment (Fig 1.2.A-(3)) (Gilula and Satir, 1972; Nachury et al., 2010). In the ciliary gate area, there are several multiprotein complexes, mutations in the genes that encode some of the proteins can lead to many ciliary diseases, which will be described in further detail below (Williams et al., 2011a).

Generally, cilia can be classified into either motile or immotile (often called primary cilium). In motile cilia, the axoneme in (most) motile cilia has a canonical 9 + 2 microtubule organisation i.e. nine doublet MTs arranged symmetrically around a pair of single MTs called the central pair (Fig 1.1.B) (Satir and Christensen, 2007). The nexin links link MT doublets together, and the central pair is connected to the MT outer doublets by radial spokes (Pigino and Ishikawa, 2012). Both inner and outer dynein arms are located on the A tubules of the MT outer doublets, which allow the dynein motor domains to contact the neighbouring B-tubule to generate ciliary movement (King, 2016).

In contrast, primary cilia have a 9+0 arrangement, in which the axoneme lacks central pair microtubules as well as radial spokes, nexin links and inner- and outer-arm dynein complexes. Although the 9+2 and the 9+0 arrangement are the two most commonly found MT structures, a wide variety of other arrangements exist for both motile and non-motile cilia. For example, the cilia on the notochordal plate of the rabbit embryo have 9+4 MTs arrangement (Feistel and Blum, 2006), and more extreme case such as gregarine apicomplexan parasites *Lecudina tuzetae*, which have 6+0 arrangement of their male gametes (Kuriyama et al., 2005). Although, the majority of motile cilia consist of a 9+2 arrangement, and primary cilia have 9+0 configurations, motile cilia can also be 9+0, or primary cilia can have a 9+2 arrangement; such as monocilia at embryonic node that generate fluid flow for left-right asymmetry (Nonaka et al., 1998), and immotile cilia in frog olfactory epithelium with 9+2 arrangement (Reese, 1965).

For many years, primary cilia were considered vestigial organelles without any real function. However, it is now recognised that primary cilia serve crucial roles, acting as a sensory antenna to allow cells to sense their environments, coordinate signalling pathways, and (in multicellular organisms) maintain tissue homeostasis (Fliegauf et al., 2007). Indeed, it is now evident that ciliary dysfunction underpins many devastating human genetic diseases, collectively termed ciliopathies (Adams et al., 2008; Berbari et al., 2009; Bloodgood, 2010; Christensen et al., 2008).

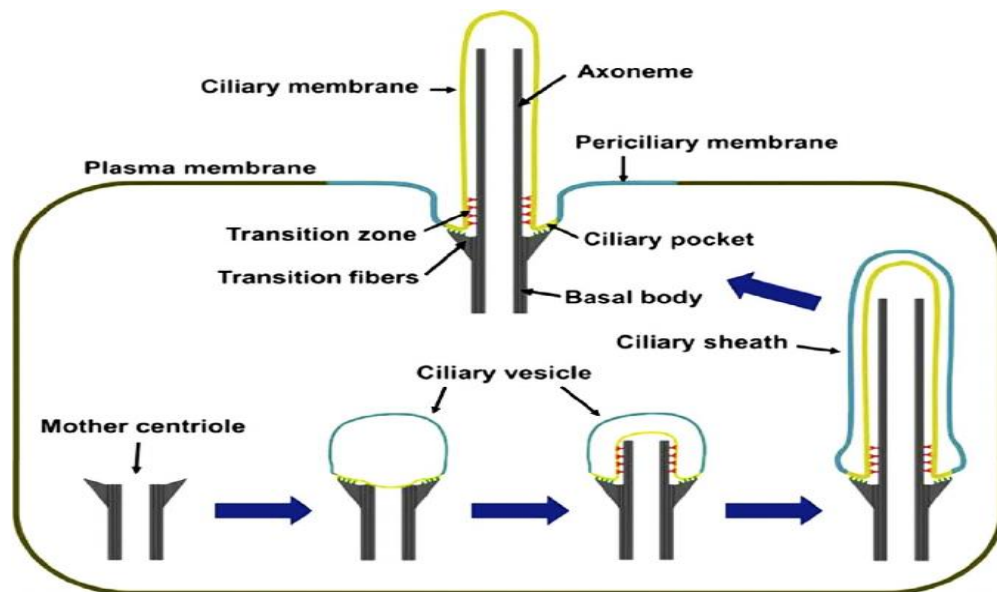




**Figure 1.2. (A)** The general structure and the cross section of a cilium at the level of the basal body, transition fibres, transition zone. **(B)** The cross section axoneme microtubule in a typical motile cilium with major structures. Adapted from (Ginger et al., 2008; Szymanska and Johnson, 2012) with permission.

### **1.1.3 Cilium biogenesis**

Early studies in mammalian cells indicated that ciliary/flagellar assembly proceeds through distinct and ordered steps (Sorokin, 1962; Sorokin, 1968) (see Fig 1.3), which were later confirmed using time-lapse microscopy (Westlake et al., 2011). The initial stage requires Golgi derived cytoplasmic vesicles, called distal appendage vesicles (DAVs), which accumulate at the distal end of the mother centriole (Kobayashi et al., 2014; Lu et al., 2015; Schmidt et al., 2012). Vesicular fusion then produces a membranous cap (also called ciliary vesicle) on the distal tip of the mother centriole (primary ciliary vesicle) (Sorokin, 1962), the microtubules of the centriole continue to grow underneath the cap, and subsequent vesicular trafficking keeps expanding the cap concomitant with microtubule extension. In this complex, the time of the sequence of events are still poorly understood. However, some evidence suggests the transition zone MT doublets, which are linked to the overlying membrane by Y-shaped linkers have already formed at this stage (Dean et al., 2016; Reiter et al., 2012; Wei et al., 2015). Later the ciliary membrane starts to form and the axoneme MTs are growing unsheathed in a double membrane. The newly formed cilium then docks to the plasma membrane by fusion with the ciliary sheath, establishing other ciliary compartments (Sánchez and Dynlacht, 2016).



**Fig 1.3 Ciliogenesis; how cilium/flagellum assembled.** The process starts when a mother centriole contacts a ciliary vesicle. Axoneme elongates at the tip and is constructed from proximal to distal, with the most proximal region giving rise to the transition zone. The ciliary vesicle continues to grow with the axoneme and gives rise to the ciliary sheath, finally the axoneme within the ciliary membrane project out from the surface and the outer sheath transforms into the periciliary membrane and fuse with the plasma membrane. Taken from (Garcia-Gonzalo and Reiter, 2012) with permission.

### 1.1.4 Ciliary signalling

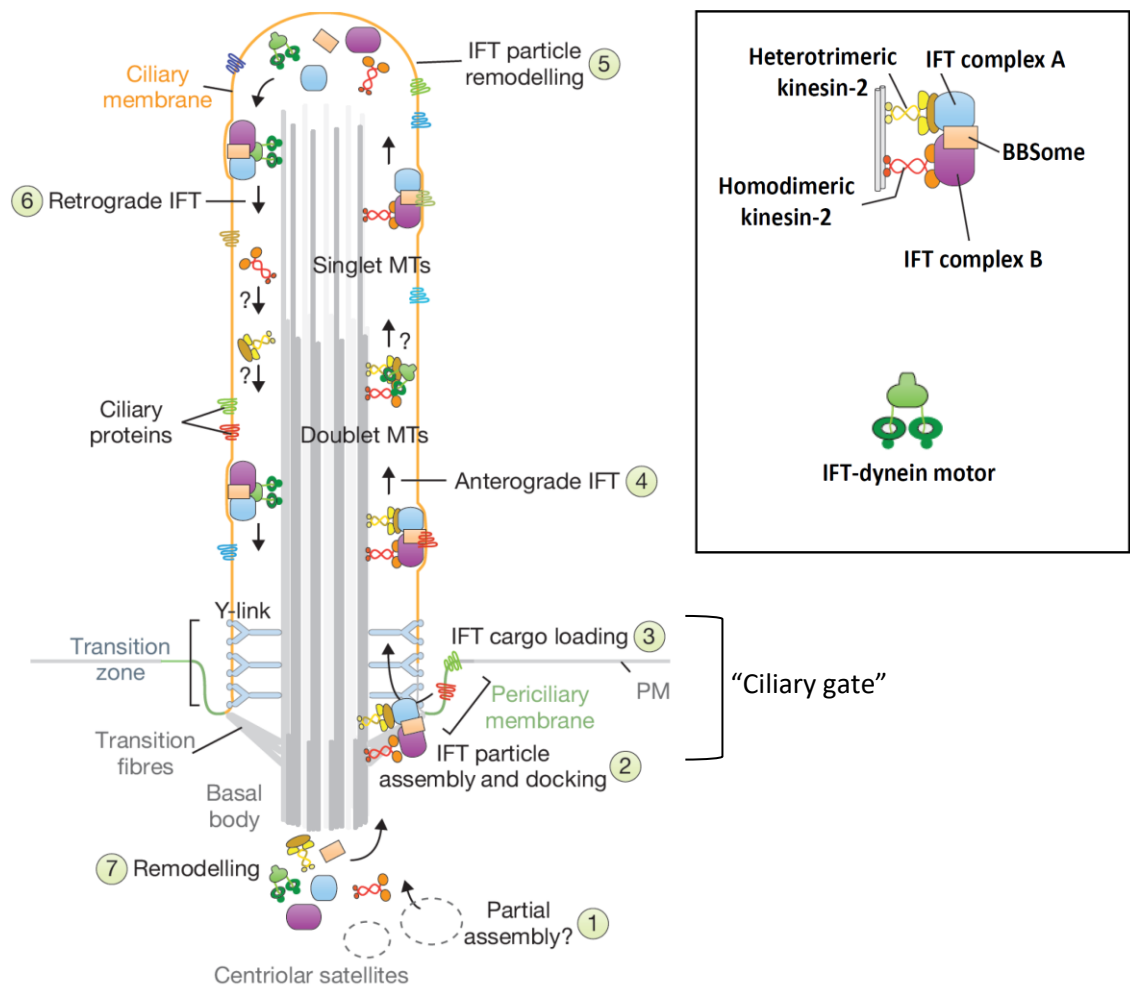
Cilia and flagella not only function as organelles of motility, but also detect environmental or intercellular stimuli. The signalling functions of cilia require regulated import and export of ciliary components, which allows those signalling molecules into and out of cilia in a semi-closed system. In recent years, a number of core signalling pathways have been described that depend on intact cilia including Hedgehog (Hh) and Wnt signalling pathways which are important cilium-associated signalling pathways in mammals; (Garcia-Gonzalo and Reiter, 2012).

## 1.1.5 IFT machinery

### 1.1.5.1 IFT motors

As mentioned earlier, ciliary assembly and maintenance relies on an evolutionary conserved mechanism known as IFT. The first description of IFT was over two decades ago in the green algae *Chlamydomonas reinhardtii*, where IFT particles were observed to move bi-directionally along the flagellum (Kozminski et al., 1993) (Fig 1.4). There are two distinct motor protein superfamilies — kinesins and dyneins that power the IFT-trains and transport IFT complex particles A and B. Each of the motor proteins only travels in one direction; anterograde IFT (base to tip) movement is facilitated by kinesin-2, and retrograde IFT (tip to base) is driven by cytoplasmic dynein 2 (Pedersen and Rosenbaum, 2008; Rosenbaum and Witman, 2002). It has been suggested that in mammals, the stage at which IFT engages and begins to assemble is shortly after the ciliary necklace has formed, and when the distal appendages bind to the membrane before axoneme assembly (Rohatgi and Snell, 2010; Sorokin, 1962). There are two types of kinesin-2 motors; heterotrimeric (consisting of two heterotrimeric kinesin-1 motor subunits and KAP (kinesin-associated protein)) and a homodimeric kinesin that together contributes to anterograde IFT. Both types of kinesin-2 motors are essential for assembly and maintenance of cilia/flagella in most of ciliated organisms (Ishikawa and Marshall, 2011). For example, *C. reinhardtii* (*fla10*) null mutants (one of the Kinesin-2 motor subunits), results in defective flagellum assembly. Moreover, *fla10* mutants appeared to be temperature-sensitive, as they failed to form flagella at the restrictive temperature (32°C), but retained flagella at a lower (21°C) temperature (Kozminski et al., 1995; Walther et al., 1994). In *C. elegans* cilia, loss of KAP subunit function on an *osm-3* (homodimeric kinesin-2 motor) mutant background resulted in the axoneme being completely absent (Evans et al., 2006a; Snow et al., 2004). The results suggest that KAP stabilises the heterodimeric coiled-coil structure formed between motor subunits in order to target the motors to the cilia and along the axoneme (Mueller et al., 2005).

Cytoplasmic Dynein 2 is part of a multiprotein complex that drives retrograde transport, returning the IFT train from the ciliary tip back to the cell body; the complex consists of cytoplasmic dynein-2 heavy chains, intermediate chain, light intermediate chain and light chain (Ishikawa and Marshall, 2011). Early studies in *C. reinhardtii*, showed that deletion of the dynein light chain (LC8) gene resulted in cells producing short and paralysed flagella that progressively shortened (Pazour et al., 1998). In these cells, anterograde IFT appeared to be normal, but retrograde IFT was decreased or absent. Large numbers of IFT particles accumulated at the tip of the flagella, suggesting that loss of cytoplasmic dynein (LC8) affected retrograde transport (Pazour et al., 1998). Later work by Yang and co-workers, indicated that LC8 is a component of inner and outer dynein arms, and is involved in stabilising outer and inner dynein arms (Yang et al., 2009). Therefore, it was not clear which motor actually causes defective retrograde IFT. Deletion of the gene encoding cytoplasmic dynein 1b heavy chain (DHC1) in *C. reinhardtii* also resulted in very short flagella and the redistribution of raft proteins from their original peri-basal body region to the flagella. However, there was no evidence of any effect on the growth and Golgi apparatus (Pazour et al., 1999). The role of the dynein light intermediate chain (D1bLIC) was also studied in *Chlamydomonas* (Reck et al., 2016), their data illustrated that knockdown of D1bLIC de-stabilised the DHC1b complex, as well as reduced both the frequency and velocity of retrograde IFT, but not completed abolishing transport. Flagellar assembly phenotypes depend on their sensitivity to the amount of active dynein 1b motor in the cell (Reck et al., 2016). In *C. elegans*, mutations in the CHE-3 gene (homologues of DHC1b) gave similar phenotypes seen in DHC1b mutant (Wicks et al., 2000). In mice model, null mutations also result in short, stumpy (<1  $\mu$ m) cilia and the IFT particles became distended (Huangfu and Anderson, 2005).



**Fig 1.4 The assembly and trafficking of the primary cilium by intraflagellar transport (IFT).**

(1) Golgi-derived centriolar vesicle (CV) localises to the distal end of the centrosomal mother centriole, and migrate toward the plasma membrane. (2) A complete formation of IFT anterograde particles (IFT complex A and complex B, BBSome and the kinesin motors) near or at the basal body transition fibre. (3) Ciliary cargo proteins are loaded onto the IFT particles. (4) Cargo is transported from the base to the tip of the cilium by anterograde motor(s). (5) When cilium grows at certain length, IFT particle remodelling occurs. (6) The retrograde dynein motors will transport ciliary cargo from tip to the base to complete the IFT cycle. (7) And the whole process will start again. The question marks indicate that unknown questions of how dynein and kinesin machineries move in the cilium. Adapted from (Sung and Leroux, 2013) with permission.

### 1.1.5.2 IFT particles

IFT particles were first isolated in *C. reinhardtii* and can be divided into two large biochemically distinct protein complexes; named IFT complex A and complex B (Cole et al., 1998; Piperno and Mead, 1997). To date, there are at least 22 proteins that have been identified. IFT complex A consists of six proteins; and complex B sixteen proteins; most of which are conserved across ciliated/flagellated organisms (see table 1.1) (Fan et al., 2010; Ishikawa and Marshall, 2011; Piperno et al., 1998; Wang et al., 2009). During anterograde IFT, both inactive dynein-IFT complex A and active kinesin-IFT B complex travel together along the cilium. Dynein-IFT complex A remains inactive until anterograde movement is complete, i.e. at the distal tip of the elongating axoneme. At this point motor activity swaps, with the dynein motor becoming active to drive retrograde movement of IFT particles (Ishikawa and Marshall, 2011). There are phenotypic differences in mutants lacking IFT complex A and B components. Null mutants of most of the IFT complex B proteins result in a more severe ciliogenesis defect. For example, IFT52 along with other 5 proteins (IFT 88, 81, 74, 46 and 27) together were originally identified as core IFT complex B protein in the green algae (Cole, 2003), and interaction organisation was characterised by Lucker et al. (Lucker et al., 2005). Their data indicated that IFT52 locates at the centre of the IFT complex B and interacts with at least five other proteins. Depletion of IFT52 leads to *C. reinhardtii* being unable to form flagella (Brazelton et al., 2001; Deane et al., 2001), and also strong destabilisation of IFT complex B subunits (except IFT27/25) (Richey and Qin, 2012). Other studies on a *C. reinhardtii* IFT complex B mutant (*ift74-2*) also showed similar destabilisation of complex B proteins (IFT20, 57, 46 and 81). Interestingly, the level of IFT139 (complex A protein) also showed a slight increase (Brown et al., 2015). Moreover, similarly increased level of IFT complex A proteins have been observed in a IFT70 knockdown background (Fan et al., 2010). Members of IFT complex B have also been shown to be involved in transport and regulating signalling pathways. For instance, IFT25 knockout mice did not show any defects in ciliogenesis or structural development, however without IFT25, membrane proteins includes Protein patched homolog 1 (Ptch1), GLI family zinc finger protein 2 (Gli2) and Smoothened (Smo) fail to activate Hedgehog pathway upon stimulation (Keady et al.,

2012), and similar defective Hh signalling also seen in null mutant of IFT27 background, despite IFT25 presence (Eguether et al., 2014).

In contrast, analysis of IFT complex A mutants showed the complex is not essential for the assembly of cilia. For example, *C. reinhardtii fla15* (IFT144) mutants and *fla17* (IFT139) mutants were still able to produce flagella, but were swollen and accumulated IFT particles (Iomini et al., 2001; Iomini et al., 2009; Piperno et al., 1998). In the mouse model, similar phenotypes have been observed in null mutants for *Ift139a/Ttc21b* and *Ift122*, where normal length cilia can be formed but the rate of retrograde IFT was reduced and cilia tips were distended (Cortellino et al., 2009; Tran et al., 2008). However, null mutations in *Ift121/Wdr35* result in a failure to produce cilia (or production of short cilia) in both humans and mice (Mill et al., 2011).

Mass spectrometry based analysis by Boldt and co-workers did not detect specific direct interaction between IFT complexes A and B (Boldt et al., 2016). However, a study in *C. reinhardtii* suggested that IFT74 (a complex B component) could be a linker between the two complexes, as truncated IFT74 (lacking 1–196 aa, including a region of coiled-coil) disturbed the interaction of IFT A and B at the base of the flagellum. This resulted in IFT complex A particles accumulating at the base of the flagellum, and IFT complex B proteins were stuck at the tip (Brown et al., 2015).

In addition to IFT complexes, there is another important protein complex called the BBosome, which contains eight conserved Bardet–Biedl syndrome proteins; BBS1, BBS2, BBS4, BBS5, BBS7, BBS8, BBS9 and BBS10 (Loktev et al., 2008; Nachury et al., 2007). These proteins move along the ciliary axoneme at the same rate as IFT particles, which have given rise to the idea of that the BBosome may act as an adaptor between the IFT complexes and IFT cargo (Mykytyn et al., 2004; Nachury et al., 2007; Ou et al., 2005). Work carried out using *C. elegans*, *C. reinhardtii* and human cells has suggested that some BBS proteins are also involved in vesicle transport from the Golgi to the basal body and cilia. A study in mice indicated that the WDR19/IFT144 (the homologue of DYF-2) and BBS1 (the homologue of BBS-1) show strong interaction, and normal cilia form in hypomorphic mutations of *dyf-2* and *bbs-1*, but



IFT complex B proteins turnaround at the ciliary tip is severely abrogated (i.e an accumulation of IFT-B components at the ciliary tip) (Liem et al., 2012; Wei et al., 2012). It remains unclear why separation of IFT complexes leads to ciliary defects. This might be due to the fact that BBSome proteins are somehow associated with the interaction between both IFT complexes and cargo transport (Blacque et al., 2004; Ou et al., 2005; Sung and Leroux, 2013). In general, BBSome proteins are not required for ciliogenesis (except BBS1 and BBS5, which fail to form cilia in siRNA treated cells) (Loktev et al., 2008; Mykytyn et al., 2004; Nachury et al., 2007; Williams et al., 2014). However, IFT machinery is required in almost all organisms for cilium/flagellum formation. This could indicate that the BBSome only transports a specific set of transmembrane proteins to cilia while the IFT complexes are likely to be required for all transport processes inside cilia (Jin et al., 2010; Lechtreck et al., 2013).

Complex	General Name	<i>Chlamydomonas reinhardtii</i>	<i>Caenorhabditis elegans</i>	<i>Homo sapiens</i>	<i>Trypanosoma brucei</i>
IFT complex A	IFT144	IFT144	DYF-2	WDR19	IFT144
	IFT140	IFT140	CHE-11	IFT140	IFT140
	IFT139	IFT139	ZK328.7	THM1 TTC21B	IFT139
	IFT122	IFT122/FAP80	DAF-10	IFT122 WDR10	IFT122
	IFT121	IFT121	IFTA-1	WDR35	PIFTD4
	IFT43	IFT43	IFT-43	IFT43 C14orf179	IFT43
IFT complex B	IFT88	IFT88	OSM-5	IFT88	IFT88
	IFT81	IFT81	IFT-81	IFT81	IFT81
	IFT74/IFT72	IFT74/IFT72	IFT-74	IFT74/IFT72	IFT74/IFT72
	IFT70	IFT70/FAP259	DYF-1	TCC30A TTC30B	PIFTB2
	IFT56	IFT56/DYF-13	DYF-13	-	PIFTB3
	IFT52	IFT52/BLD1	OSM-6	IFT52 NGD5	IFT52
	IFT46	IFT46	DYF-6	IFT46 C11orf60	IFT46
	IFT27	IFT27	Absent	IFT27 RABL4	IFT27
	IFT25	IFT25/FAP232	Absent	IFT25 HSPB11	IFT25
	IFT22	IFT22/FAP9	IFTA-2	RABL5	IFT22
	IFT172	IFT172	OSM-1	IFT172	IFT172
	IFT80	IFT80	CHE-2	IFT80 WDR56	IFT80
	IFT57	IFT57	CHE-13	IFT57	IFT57
	IFT54	IFT54/FAP116	CHE-11	IFT54 TRAF3IP1 MIPT3	IFT54
	IFT38	IFT38/FAP22	DYF-3	IFT38	PIFTA1
	IFT20	IFT20	IFT-20	IFT20	IFT20

**Table 1.1 Components of the IFT proteins in different organisms.** CHE, chemotaxis abnormal; DYF, abnormal dye-filling; WDR, WD repeat-containing; TTC, tetratricopeptide repeat; THM1, TTC-containing Hedgehog modulator; DAF, abnormal dauer formation; OSM, osmotic avoidance abnormal; PIFT, putative involved in IFT; ZK328.7, Tetratricopeptide repeat protein 21 homolog; FAP, flagellar-associated protein; TRAF3IP1, TRAF3-interacting protein 1; IFT, intraflagellar transport; MIPT3, microtubule-interacting protein associated with TRAF3; IFTA, IFT-associated; KAP, kinesin-associated protein; Table is adapted from (Ishikawa and Marshall, 2011; Taschner and Lorentzen, 2016).

## 1.2 The base of the cilium

### 1.2.1 The ciliary gate

In a cilium, the junction between the basal body and the axoneme is called the ciliary gate, and can be divided into two structurally distinct sub-regions; **(1) the transitional fibres (TFs)** (also called the alar sheets) (Anderson, 1972), which serve as the docking site for both the cargo that requires import into the flagellum and IFT particle proteins. There are nine TFs, and each one emerges from the distal end of the C-tubule of the basal body, and terminates at an electron dense knob at the most proximal end of ciliary membrane (Lu et al., 2015; Sorokin, 1962). TEM studies suggest that the space between each TF is less than 60 nm in diameter, which might allow a ribosome to pass through but not a vesicle (Nachury et al., 2010). So far, at least five proteins have been identified that localise to the TFs. Centrosomal protein of 164 kDa (Cep164) is a key component of the TFs, and has an important role in ciliogenesis, as depletion of Cep164 can lead to ciliary assembly defects and blocks DAV formation (Graser et al., 2007). When Cep 164 was absent, IFT88 (IFT B complex component) failed to recruit to the basal body and ciliogenesis was prevented (Schmidt et al., 2012; Čajánek and Nigg, 2014). By using super resolution microscopy techniques such as PALM and STORM analysis, another centrosomal protein, Cep89/Cap123/CCDC123, has also been localised to the centriolar appendages, and forms ring-like structures with a diameter of 500 nm (Sillibourne et al., 2011). Three other novel DAP proteins (Cep83/Ccdc41, SCLT1 and FBF1) were identified by Tanos and colleagues using quantitative centrosome proteomics and super resolution microscopy (Tanos et al., 2013), they revealed a hierarchy of DAP assembly. Both SCLT1 and Cep89 are required Cep83 to the centrioles. The centriolar localisation of FBF1 and CEP164 are not dependent on each other or Cep89, but is mediated by SCLT1. All three proteins have been shown to be involved in ciliogenesis and the docking of the basal body to the plasma membrane. The consequences of the basal body being unable to dock can lead to a failure to recruit Tau-tubulin kinase 2 (TTBK2) or the release of centriolar coiled-coil protein of 110 kDa (Ccp110). *In vitro*, TTBK2 is able to phosphorylate microtubule-associated proteins (Takahashi et al., 1995), and *in vivo* it acts as a key regulator for

controlling the initiation of ciliogenesis by removing basal body capping protein Cp110 from the distal end of the mother centriole, permitting initiation of ciliogenesis (Bouskila et al., 2011; Goetz et al., 2012). However, TTBK2 is not essential for basal body formation (Goetz et al., 2012). TF components also participate in recruiting IFT particles and facilitating the passage of assembled IFT particles through the TF. For instance, Cep83 and FBL1 both bind IFT complex B protein IFT20 and IFT54 respectively, null mutation Cep83 and FBL1 failed IFT particles to recruit to the basal body and promote to them pass the ciliary gate to enter cilium, which lead to a ciliogenesis defect (Wei et al., 2013; Ye et al., 2014) .

TFs serve an important function in recruiting proteins bound for the cilium. However, exactly how TFs actively promote proteins to pass the ciliary gate is unknown. Ludington et al. (2013), proposed an avalanche-like behaviour in ciliary import, i.e. after an accumulation of IFT proteins at the cilium base for a period of time, a large number of IFT particles enter the cilium together. Whether this avalanche-like behaviour depends on the amount of IFT accumulating on the TFs is unclear.

**(2) The transition zone (TZ)** is the boundary between the basal body and the axoneme, and is an area with high molecular complexity. Within the transition zone, there is a distinctive and characteristic structure called the Y-shaped linkers (Y-linkers), which connect axonemal microtubule outer doublets to the ciliary membrane and form a unique particle arrangement called the ciliary necklace (Gilula and Satir, 1972). Although the protein components of both Y linkers and the necklace are largely uncharacterised, it is generally accepted that the TZ functions as a ciliary gatekeeper that regulating protein enter and exit the cilium (Reiter et al., 2012). For many years, transmission electron microscopy (TEM) cross-sections were used to investigate the TZ, and data suggested that it forms at the very earliest stages of ciliogenesis that precedes the recruitment of IFT particles (Omran, 2010; Ringo, 1967; Williams et al., 2011; Chih et al., 2012; Craige et al., 2010; Williams et al., 2011b).

### 1.2.2 The ciliary gate and ciliopathies

For many years, the function of cilium/flagellum was thought to only involve motility and primary cilia were long thought to be evolutionary vestigial organs. However, increasing evidence recognised that primary cilia act as multi-functional organelles associated with a series of important signaling pathways (e.g. Hedgehog (Corbit et al., 2005) and Wnt signaling pathways (Corbit et al., 2008)) during early vertebrate embryonic development and maintain postnatal tissue homeostasis (Chih et al., 2012; Gerdes et al., 2009; Goetz and Anderson, 2010; Gonçalves and Pelletier, 2017).

In humans, dysfunction or defects in motile and primary cilia are now understood to underlie a number of devastating genetic conditions, which are termed ciliopathies (Fliegauf et al., 2007; Novarino et al., 2011; Tobin and Beales, 2009; Veland et al., 2009). Ciliopathies can affect single organs, such as Polycystic kidney disease (PKD) or can develop as multi-system disorders such as Bardet–Biedl syndrome (BBS), Nephronophthisis (NPHP), Joubert Syndrome (JBTS) and Meckel-Gruber Syndrome (MKS), which have phenotypically variable and overlapping disease manifestations (Badano et al., 2006; Cardenas-Rodriguez and Badano, 2009). Common clinical features in ciliary dysfunction syndromes include polycystic kidneys, liver fibrosis, retinal degeneration, brain/nervous system defects and skeletal deformities (Fliegauf et al., 2007; Gerdes et al., 2009). Clinical data also shows that defects in cilia formation can be observed in human cancers, as key developmental signalling pathways are regulated from the primary cilium (Seeger-Nukpezah et al., 2013). For example, Hedgehog signalling activity is required for IFT proteins to be transported to and from the tip of primary cilia, when cilia formation is blocked, alteration of signalling activity can promote cancer development (Goetz and Anderson, 2010; Huangfu et al., 2003).

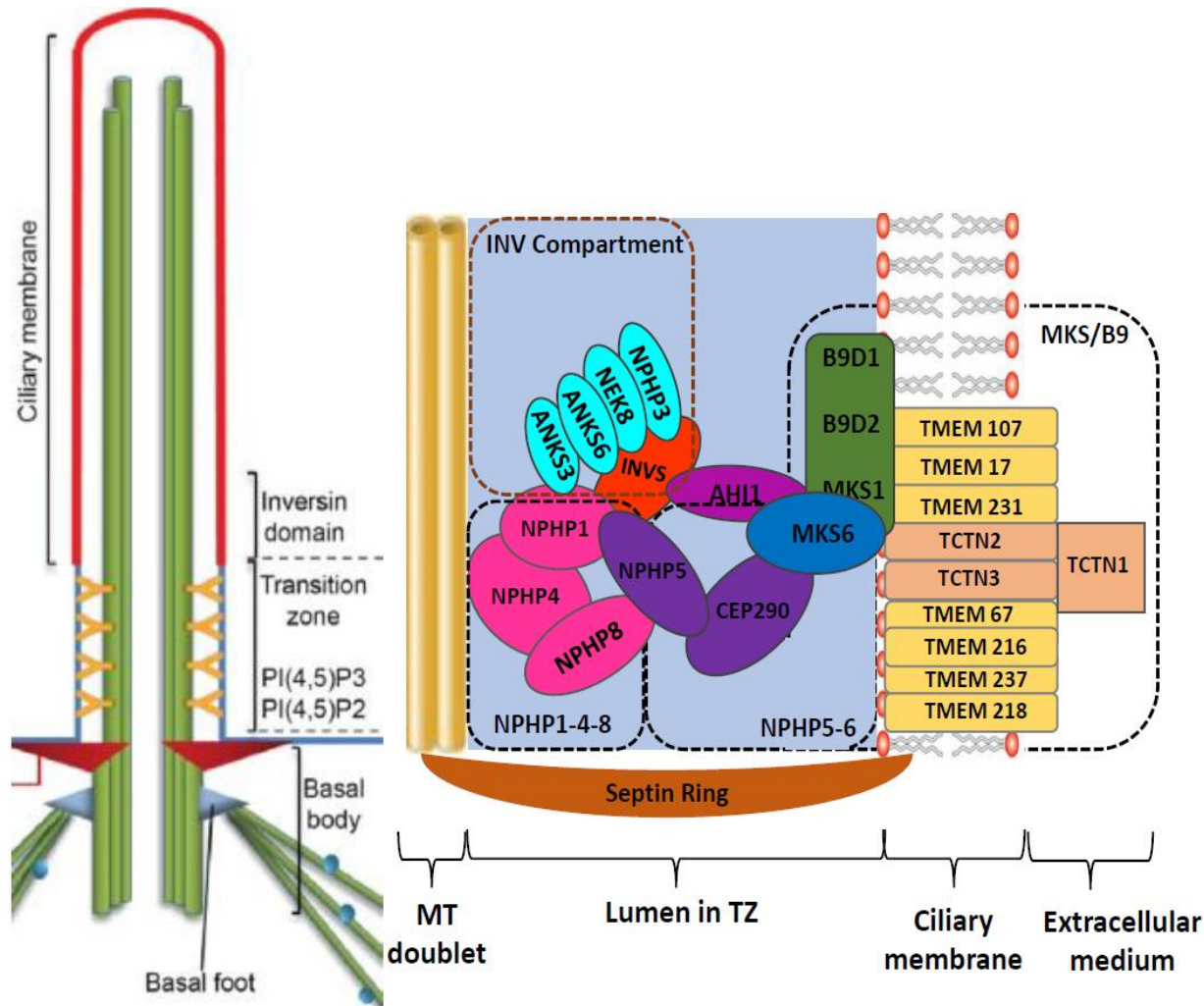
Several studies have revealed that there are large number of proteins encoded by genes mutated in NPHP, JBTS and MKS that are localised to the basal body/TZ area (Sang et al., 2011). This area acts as a barrier to control protein entry into, and exit from, the cilium; hence the term “ciliary gate” (Anand and Khanna, 2012; Bachmann-Gagescu et al., 2011; Chih et al., 2012; Czarnecki and Shah, 2012; Garcia-Gonzalo et

al., 2011; Williams et al., 2011; Garcia-Gonzalo and Reiter, 2017). Over the past few years, several research groups have investigated the composition of protein complexes at the ciliary gate area, and have provided insight into (1) how these proteins might interact; (2) how they assemble within the TZ; (3) their specific functions within the ciliary compartment; and (4) how these proteins are targeted to the ciliary gate (See Table 1.2 and Fig 1.5). A number of distinct complexes have been identified at the region of ciliary gate – these are:

(1) **NPHP1-4-8** (NPHP1/JBTS4; NPHP4 and NPHP8/RPGRIP1L)

(2) **NPHP5-6** (NPHP5/IQCB1 and CEP290/NPHP6/MKS4/TBTS5)

(3) **MKS/TCTN/B9** (MSK1; MKS2/TMEM216; MKS3/NPHP11/JBTS3; CC2D2A/MKS6/JBTS9; MKS8/TCTN3; TCTN1/JBTS13; TCTN2/MKS8; TCTN3; B9D1/MKS9; B9D2/MSK10; TMEM237/JBTS14)



**Figure 1.5 Model organization chart of four protein complexes and the possible connections**

The **NPHP1-4-8 protein complex** consist NPHP1, NPHP4 and NPHP 8/RPGRIP1L. NPHP8/RPGRIP1L localizes the TZ, but also to the BB. Evidences suggested that NPHP4 directly interact with RPGRIP1L and NPHP1, but no direct interaction between NPHP1 and NPHP1. The **MKS/B9 protein complex**, MKS1, TMEM216 and TMEM67, is consistently associated with severe ciliopathies in human. Cep290 was considered as a core candidate of **NPHP5-6 protein complex** (Sang et al., 2011), but it may also be part of MKS-like complex in another study (Garcia-Gonzalo et al., 2011). The **inversin compartment** protein candidates may serve a bridging role and contribute to multiple distinct roles. Reproduced from (Garcia-Gonzalo and Reiter, 2017; Gonçalves and Pelletier, 2017) with permission.

**Table 1.2 List of overviews of human ciliopathic-associated genes encode proteins**

Protein Complex	Gene Name	Ciliopathy names			Localisation	Protein Domain	Selected References
		NPHP	MKS	JBTS			
<b>NPHP 1-4-8</b>	RPGRIP1L	NPHP8	MKS5	JBTS7	TZ, BB	CC, C2	(Garcia-Gonzalo et al., 2011; Huang et al., 2011; Sang et al., 2011; Williams et al., 2011)
	NPHP1	NPHP4			TZ	CC, C2, SH3	(Sang et al., 2011; Williams et al., 2011)
	NPHP4	NPHP1		JBTS4	TZ	C2, MSP	(Garcia-Gonzalo et al., 2011; Huang et al., 2011; Sang et al., 2011; Williams et al., 2011)
<b>NPHP 5-6</b>	CEP290	NPHP6	MKS4	JBTS5	TZ, BB	CC	(Garcia-Gonzalo et al., 2011; Huang et al., 2011; Sang et al., 2011; Williams et al., 2011)
	IQCB1	NPHP5			BB	IQ	(Williams et al., 2011; Garcia-Gonzalo et al. ; Sang et al., 2011)
<b>MKS/TCTN/B9</b>	B9D1		MKS9		TZ, BB	B9	(Williams et al., 2011; Sang et al., 2011; Chih et al., 2012; Dowdle et al., 2011)
	B9D2		MKS10		TZ	B9	(Williams et al., 2011; Sang et al., 2011; Dowdle et al., 2011)
	TCTN1			JBTS13	TZ	TCTN,	(Garcia-Gonzalo et al., 2011)
	TCTN2		MKS8		TZ, AX	TCTN, TM	(Garcia-Gonzalo et al., 2011; Sang et al., 2011)
	TCTN3				TZ, AX	TCTN, TM	(Garcia-Gonzalo et al., 2011)
	CC2D2A		MKS6	JBTS9	TZ	CC, C2	(Garcia-Gonzalo et al., 2011; Huang et al., 2011; Sang et al., 2011; Williams et al., 2011)
	MKS1		MKS1		TZ	B9	(Williams et al., 2011; Garcia-Gonzalo et al., 2011; Sang et al., 2011; Goetz et al., 2017)
	TMEM17				TZ	TM	(Chih et al., 2012; Li et al., 2016)



	TMEM107				TZ	TM	(Lambacher et al., 2016)
	TMEM213		MKS11	JBTS20	TZ	TM	(Chih et al., 2012)
	TMEM216		MKS2	JBTS2	TZ, BB	TM	(Garcia-Gonzalo et al., 2011; Huang et al., 2011; Valente et al., 2010)
	TMEM67		MKS3	JBTS6	TZ, AX	TM	(Iannicelli et al., 2010; Williams et al., 2011; Garcia-Gonzalo et al. 2011)
	TMEM237			JBTS14	TZ	TM	(Huang et al., 2011)
<b>Inversin Compartment</b>	IVN	NPHP2			TZ, BB, AX	Ankyrin, IQ	(Yokoyama et al., 1993; Otto et al., 2003; Simons et al., 2005; Shiba et al., 2010)
	NPHP3	NPHP3	MKS7		AX	ATPase, TPR	(Olbrich et al., 2003; Shiba et al., 2010)
	NEK8	NPHP9			AX	Ser/Thr, RCC1	(Zhang et al., 2002; Attanasio et al., 2007)
	ANKS3				AX	Ankyrin, SAM	(Delestré et al., 2015; Leettola et al., 2014; Yakulov et al., 2015)
	ANKS6	NPHP16			AX	Ankyrin, SAM	(Czarnecki et al., 2015; Delestré et al., 2015; Hoff et al., 2013; Leettola et al., 2014)

TZ, Transition zone; BB, basal body; AX, axonemal compartment; CC, coiled coil domain; TM, transmembrane; SH3, Src homology 3; TPR, tetratricopeptide repeat; WD40, WD40- or b-transducin domain; RCC1, regulator of chromosome condensation 1-like; C2, B9, membrane-association domain; C2H2, zinc finger-like domain; IQ, IQ calmodulin-binding domain.

### 1.2.2.1 NPHP1-4-8 protein complex

NPHP1-4-8 module was first identified by Sang et al., (Sang et al., 2011) and shown to consist of (i) NPHP1/JBTS4 protein - previously suggested to be involved in regulating IFT particle entry into the cilium (Jiang et al., 2009); (ii) NPHP4- mutations in the gene encoding *NPHP4* lead to nephronophthisis and blindness in humans (Hoefele et al., 2005; Otto et al., 2002), and (iii) NPHP8 (also called RPGRIP1L/MSK5/JBTS7) (Arts et al., 2007; Delous et al., 2007).

Several studies have investigated the function of NPHP4 in different models, and suggest that when NPHP4 is absent, it can lead to phenotypes indicating a ciliary defect. For example, in zebrafish, morpholino-mediated knockdown of NPHP results in typical ciliary-defect phenotypes such as abnormal body curvature and pronephric cysts (Burcklé et al., 2011). In mice, NPHP4 is required for formation and maintenance of photoreceptor cells and sperm development (Won et al., 2011). Later studies in *C. reinhardtii* also demonstrated that NPHP4 is involved in regulating both membrane and soluble protein composition at the flagellum TZ (Awata et al., 2014).

A study by Sang et al., (Sang et al., 2011) showed that the NPHP1-4-8 complex localises to the TZ and *in vitro* studies revealed that NPHP4 directly binds both NPHP1 and NPHP8; acting as a bridge between them, but that NPHP1 and NPHP8 do not bind to each other. Based on their *in vitro* model, cilium formation is not dependent upon NPHP1, NPHP4 or NPHP8. However, when a 3D spheroid culture model involving IMCD3 cells was analysed (which represents the architecture of the kidney collecting duct with a clear lumen, apical cilia, defined tight junctions, and clear basolateral structure) the depletion of NPHP1, 4 and 8 resulted in cells developing an irregular lumen and there was loss of tight junction formation (Sang et al., 2011). All three NPHP1-4-8 proteins contain C2 domains, which could mediate interaction with phospholipids at cell-cell junctions or at the ciliary base. In mice, studies have shown that both NPHP1 and NPHP4 mutations only lead to mild cilia formation defects, in sperm flagella and cilia in photoreceptor cells (Jiang et al., 2009; Jiang et al., 2008).

Williams and colleagues (Williams et al., 2011) investigated NPHP1-4-8 and MKS protein complexes (for more details see section 1.2.2.2) in ciliated sensory neurons in *C. elegans*, and reported that NPHP8/RPGRIP1L/MKS5 was the key protein linking these two protein complexes. They proposed that NPHP8 is at the top of the hierarchical network of these protein modules, When NPHP8 gene was deleted it resulted in defective cilia formation, and also affected the recruitment of NPHP1 and NPHP4 to the TZ; it also leads to all MKS proteins failing to localise to the TZ. In *C. elegans*, disruption of proteins from either the NPHP or MKS complex do not have noticeable effects on ciliogenesis, but disruption of proteins from both the MKS and NPHP complexes results in severe phenotypes. For example, perturbation of both MKS5 and MKS6 (MKS protein complex) and NPHP4 (NPHP complex) can abrogate ciliogenesis in most cilia. Transmission electron microscopy reveals that these double mutants lack the TZ and Y linkers, which leads to membrane detachment from the ciliary axoneme (Huang et al., 2011; Jensen et al., 2015; Williams et al., 2011). However, NPHP8 is the exception, as mutations in NPHP8 in mice are embryonic lethal which is similar to the consequences of double mutants in the mouse model (affecting MKS and NPHP complexes components (Garcia-Gonzalo et al., 2011; Sang et al., 2011). This may also explain why mutations in the NPHP8 gene are present in patients with NPHP, MKS and JBTS. However, NPHP-1 or NPHP-4 are involved in less severe ciliopathies, because they do not directly interact with proteins in other complexes (Czarnecki and Shah, 2012).

Both the MKS and NPHP protein modules are required for the early stage of the ciliogenesis pathway, before the IFT-dependent axoneme extension, TZ membrane docking and TZ formation, since when single or double mutants of the MKS/NPHP complex were generated IFT is unaffected. Additionally, TZ localisation of the MKS/NPHP module proteins were unaffected in IFT mutants (*osm-5*, *che-11*) and BBS mutants (*bbs7* and *bbs8*). The results might indicate that MKS and NPHP protein complexes must be established before IFT becomes active. Together, these data suggested that NPHP1-4-8 complex is involved in apical and polarized cells organizations and participated in cell morphogenesis and architecture (Sang et al., 2011).

### 1.2.2.2 The MKS/B9 complex

The MKS/B9 protein complex identified by several studies contains three membrane-targeting B9-containing proteins (MKS1; MKS9/B9D1 and MKS10/B9D2), a coiled-coil protein MKS6/CC2D2A, transmembrane proteins TMEM67, TMEM216, TMEM17, TMEM231, TMEM107 and possibly TMEM237 and TMEM 218, as well as three Tectonic proteins (Tctn1-3) (Chih et al., 2012; Garcia-Gonzalo et al., 2011; Sang et al., 2011; Williams et al., 2011).

As with other TZ protein complexes, studies have revealed that mutations in single MKS genes (e.g. MKS6, MKSR-1, MKSR-2, MKS1, MKS3, TMEM107, TMEM237 and TMEM218) resulted in mild or no obvious defects in cilia formation; nor did they impair IFT. However, mutations in combined MKS genes and NPHP genes appeared more severe, and resulted in defective ciliogenesis and ciliary function (Huang et al., 2011; Li et al., 2016; Williams et al., 2011; Williams et al., 2008).

In mammals, knocking-out different genes coding for different MKS components can lead to tissue-specific defect in ciliogenesis. Tectonic 1 (TCTN1), (which is involved in modulating Hh signalling pathway) is required for ciliogenesis in a tissue-dependent manner, for example in TCTN1 null mutants the majority of the cilia in neural tubes appear to be shortened and bulbous, and the number of cilia is significantly reduced compared to controls. In TCTN1 null mutants, basal bodies dock at the cellular plasma membrane, but the axoneme fails to extend. In contrast, in limb bud mesenchym cells (TCTN1<sup>-/-</sup>), while the number of cilia were reduced, no obvious defects could be observed in the cilia that were formed (Garcia-Gonzalo et al., 2011).

Later studies in *C. elegans* demonstrated TCTN1 interacted with BBS1, and MKS1 interacts with BBS4 (Goetz et al., 2017; Yee et al., 2015). Deletion of TCTN1 alone, or together with other MKS complex components, did not abolish ciliogenesis. However, when TCTN1 depletion were combined with either NPHP-1 or NPHP-4 mutants, ciliary structure and ciliogenesis was compromised, possibly due to deregulation of the Hh

signalling pathway. In addition, when limb bud cells are missing both TCTN1 and BBS1, cilia formation is completely blocked (Yee et al., 2015). The study carried out by Goetz and co-workers showed strong genetic interactions between MKS1 and IFT72, as well as dynein subunit DYNC2H1 (IFT retrograde motor) (Goetz et al., 2017). Double mutants of MKS1 and either IFT72 or DYNC2H1 led to a reduced rate of ciliary protein trafficking, and a failure of cilium formation. Collectively, this data suggests that the MKS complex may cooperate with IFT to mediate axoneme assembly, and that the BBSome may have functional overlap with MKS modules in facilitating IFT dependent ciliogenesis.

Proteomic analysis of mammalian COS1 cells suggested that TCTN1 strongly interacts with TCTN2 and TCTN3; indicating that all three TCTN proteins are within the same protein complex. TCTN1 also interacts with MKS1, CC2D2A/MKS6 and B9D1, but only interacts with TMEM67 and TMEM216 under certain circumstances. For example, TMEM67 and TMEM216 can only immunoprecipitate with TCTN1-LAP and TCTN1-V5 respectively in the lysates of COS1 cells. CEP290, a possible component of the Y-shaped linkers at TZ, has also been identified as an interactor with TCTN1 by mass spectrometry and chromatography but not co-immunoprecipitation. It might have suggested that CEP290, TMEM67 and TMEM216 are peripheral components of the Tectonic complex (Garcia-Gonzalo et al., 2011). TCTN1 is required for the recruitment of both MKS1 and TMEM67 proteins to the TZ. Mutations in MKS complex components, such as TCTN2, MKS6 and TMEM67, lead to tissue-specific defects in cilia formation, protein trafficking and membrane composition (Garcia-Gonzalo et al., 2011).

A small G protein – Arl13B has been shown to be important in the trafficking of ciliary proteins and structural integrity of the cilium (Li et al., 2012). A recent study carried out by Gotthardt and co-workers identified that Arl13B actually acts as a GEF to activate Arl3 in cilia, and the GEF activity relies on its G-protein domain and C-terminal helix (Gotthardt et al., 2015). Previous studies have shown that mutations in the gene that encodes Arl13B can cause Joubert syndrome in humans (Cantagrel et al., 2008). Studies in *C. elegans* showed that ARL13B is associated with IFT complex B, via

interaction with IFT46 and IFT74, and that MKS/NPHP modules regulate the ARL13 diffusion barrier at the TZ (Cevik et al., 2010; Cevik et al., 2013). In mammal, ARL13B levels was markedly reduced in *Tctn1*<sup>-/-</sup>, *Tctn2*<sup>-/-</sup> cilia. The role of TCTN1 is more likely to involve promoting ciliary translocation of Arl13B and Smo signalling rather than limb bud ciliogenesis (Casparly et al., 2007; Corbit et al., 2005; Garcia-Gonzalo et al., 2011).

Huang et al., 2011 identified TMEM237/JBTS14 and showed that mutations in the gene encoding this protein can lead to JBTS related disorders (Huang et al., 2011). Studies in three models (mammalian IMCD3 cells, *D. rerio* and *C. elegans*) confirmed that TMEM237 is a TZ protein. TMEM237 functionally interacts with other ciliary proteins, such as B9D1/MKS9 and B9D2/MKS10, NPHP4, RPGRIP1L/NPHP-8/MKS5 and TMEM216/MKS2/TBTS2 at the TZ. Studies reveal that RPGRIP1L/NPHP-8 is required for TMEM237 to be recruited to the TZ in both mammalian and *C. elegans* cells, suggesting that RPGRIP1L acts as a connector between ciliary protein complexes (Huang et al., 2011).

TMEM231 and TMEM17 are transmembrane proteins, which might indicate that the MSK complex is attached to a membrane. In cell membrane fractions of IMCD3 and murine embryonic fibroblasts (MEFs) other MKS proteins, such as B9D1, B9D2, MKS1 and TCTN1 can also be detected, which might suggest that they are connected to the membrane via TMEM231 and TMEM17. Loss of IFT particle protein IFT88 has no effect on the localisation of TMEM231, which suggests that the TZ can assemble in the absence of IFT. More severe phenotypes are only be seen in *C. elegans* double mutants (i.e. carrying mutations in a gene encoding a protein from each of the MKS and NPHP complexes (Bialas et al., 2009; Williams et al., 2010; Williams et al., 2008). In mammals (e.g. mice), a mutation in any of the components from the MKS complex can lead to catastrophic consequences. This difference may be explained by the fact that in both worms and flies, the only place that cilia are present is in sensory neurons, and cilia serve specific sensory functions. However, in vertebrates, cilia are ubiquitous cellular organelles and TZ proteins have developed other functions/mechanisms. TMEM107

was recently identified as a component of MKS module with TZ localisation. Together with TMEM231 and TMEM216, as well as NPHP1 display a periodic localisation pattern that is highly reminiscent of ciliary necklace strands or their underlying Y-linkers (Lambacher et al., 2016). In *C. elegans*, the MKS module is proposed to assemble hierarchically (Huang et al., 2011; Roberson et al., 2015; Williams et al., 2011a). In Lambacher and co-workers *C. elegans* model, they have grouped three layers at TZ. Layer 1 contains NPHP8/MKS5, which is at the root of the hierarchy and it required for TZ localization of all MKS module components. Layer 2 or intermediate level consist MKSR-1, MKSR-2, TMEM231 and MKS2, and the localisation of the layer 2 proteins required for Layer 3 (peripheral level MKS3, JBTS14, MKS6 and TMEM17, which is not essential for other MKS protein to be localised to TZ), but not Layer 1 protein targeting. TMEM107, is required for the recruitment of some MKS module proteins (MKS1, TMEM231 and TMEM237) being to the TZ form an intermediate level. Loss of TMEM107 also perturbs ciliary composition, and together with NPHP4 plays a critical role in TZ docking, Y-linker assembly and maintenance of cilium integrity (Lambacher et al., 2016).

### **1.2.2.3 The NPHP5-6 or Cep290 protein complex**

This protein module consists of only two proteins, NPHP5 (also known as IQCB1) and CEP290 (also known as NPHP6/MKS4/JBTS5), and is localised at the centrosome/basal body (Sang et al., 2011). CEP290 was originally identified using proteomic approaches and shown to have multiple cellular localisations within mammalian cells, including the centrosomes of dividing cells, the connecting cilium of retinal photoreceptors (Chang et al., 2006) and pericentriolar satellites (Kim et al., 2008). Additionally, CEP290 also appears to have nuclear and cytoplasmic localisations (Guo et al., 2004; Sayer et al., 2006). The *C. reinhardtii*, CEP290 protein localises to the Y-shaped linkers of the TZ, and mutations in the CEP290 gene cause defects in flagellar composition, an imbalance of IFT complexes A and B in the flagellum, and lead to abnormal levels of BBS4 and PKD2 proteins, which are implicated in human ciliopathies (Craigie et al., 2010). The localisation of CEP290 is highly dynamic and rapidly turned over at the TZ,

suggesting CEP290 is associated with signaling pathways between the cilium and cell body (Craigie et al., 2010). Purification of the NPHP5-6 complex established the strong binding between these two proteins in all cell types examined (NIH 3T3 fibroblasts, IMCD3 and RPE cells), and that CEP290 is required for NPHP5 recruitment to the centrosome. Both NPHP5 and CEP290 are indispensable for ciliation in IMCD3 cells, but when either CEP290 or NPHP5 was depleted can lead to failure to form normal spheroids, suggesting their function in maintaining the integrity of centrosome/cilia in tissue organization (Sang et al., 2011). Among the MKS/NPHP proteins, only NPHP5 interacts with IFT122 (Sang et al., 2011). However, it remains to be investigated whether the NPHP protein complex is functionally linked to IFT. In addition, CEP290 also acts as an important hub between MKS and NPHP complexes via CC2D2A/MKS6 in zebrafish. Specifically, the combined loss of function of CEP290 (Knockdown) and CC2D2A (*sentinel* mutation) results in pronephric cysts formation exacerbation, revealing a genetic interaction between CC2D2A and CEP290 and implicating CC2D2A in cilium/basal body function (Gorden et al., 2008).

In *C. elegans*, CEP290 and NPHP8 are required for recruitment of TMEM218 (MKS protein) to the TZ, and together with other NPHP proteins facilitates ciliogenesis (Li et al., 2016). In the mouse model, CEP290 is required for BBSome assembly and Barbelanne and co-workers revealed CEP290 and NPHP5 physically interact with several BBSome proteins (BBS2 and BBS5), and that depletion of NPHP5 or CEP290, or expression of mutant NPHP5, causes the dissociation of BBS2 and BBS5. Furthermore, loss of either protein perturbs the integrity of BBSome, as well as the delivery of BBSome cargo to the ciliary compartment; suggesting their novel role in regulating ciliary trafficking of the BBSome and maintenance of its integrity (Barbelanne et al., 2015; Zhang et al., 2014). However, the CEP290 gene is not recognised in the genomes of many ciliated unicellular and bilaterians, which may suggest that CEP290 is not required for ciliary function in all flagellated eukaryotes, or that sequence divergence just makes this gene difficult to identify (Huang et al., 2011).



#### 1.2.2.4 Inversin compartment (IC)

An additional ciliary segment localises distal to the TZ and proximal end of axoneme, termed the “**Inversin compartment**”, which contains INV/NPHP2; NPHP3/MKS7; NEK8/NPHP9) (Shiba et al., 2010). Previously, INV mutations had been shown to lead to nephronophthisis type 2, which can be found in patients with renal cystic disease (Otto et al., 2003). Apart from INV, there are two more proteins NPHP3/MKS7 and NEK8/NPHP9 that have been identified within the INV compartment. The INV compartment is a small region that sits between the microtubule axoneme and the TZ, the exact function is yet to be understood. However, the evidence suggests INV protein is required for two other components, NPHP3 and NEK8, to localise at the INV compartment. Loss of NPHP3 and NEK8 has no effect on INV protein localisation. The data indicates that INV protein acts as a hub to interact with both NPHP3 and NEK8 in the inversin compartment, but the role of this new structure remains mysterious. It has been suggested that INV protein, along with the protein AHI1/Joubertin forms a linkage which connects with the NPHP1-4-8, NPHP5-6 and MKS protein complexes (Shiba et al., 2010). There are two more proteins identified as members of the Inversin compartment called ANKS3 and ANKS6 (Ankyrin Repeat And Sterile Alpha Motif Domain Containing protein 3 and 6). Mutations in both ANKS3 and ANKS6 can lead to autosomal dominant polycystic kidney disease (ADPKD) in mice and nephronophthisis in humans (Hoff et al., 2013; Yakulov et al., 2015). ANKS3 and ANKS6 interact directly via their SAM domains (Leettola et al., 2014), and co-localise in mouse renal cilia (Delestré et al., 2015). Yakulov and colleagues have recently shown that ANKS6 forms a protein complex containing INV, NPHP3, and NEK8. In zebrafish embryos, loss of ANKS3 causes NPHP-typical phenotypes, including ciliary defects and cyst formation. Moreover, ANKS3 also shows an association with NPHP1-4-8 protein complex via NPHP1 in multi-ciliated epidermal cells (Yakulov et al., 2015). ANKS6 has also been identified as activator and link to NEK8 to INV and NPHP3, without NEK8, ANKS6 failed to localise to the ciliary INV compartment. Mutations in ANKS6 (Strkr) can decrease its ability to interact with NEK8, and also prevent NEK8 binding to its kinase domain. On the other hand, mutant NEK8 causes a loss of kinase function, but localisation of ANKS6 is unaffected at Inversin compartment (Czarnecki et al., 2015; Hoff et al., 2013).

### 1.2.2.5 Other proteins and lipids

Apart from classic TZ proteins, there are number of other proteins and lipids which may be involved in establishing TZ complexes. These include septins, which are a large, conserved family of GTPases that form large, ordered structures such as apolar filaments, bundles and rings. They play important roles in mitosis, cell migration, and cell morphogenesis by forming scaffolds and diffusion barriers (Hu and Nelson, 2011; Hu et al., 2008). Septin proteins usually co-operate with other septins, for example in sperm cells Septin 7 is co-localised with Septin 4 at the annulus that forms at the base of the axoneme early in sperm flagellum biogenesis. Septin 4 is also part of the diffusion barrier that regulates flagellar proteins at the flagellum base (Ihara et al., 2005; Kissel et al., 2005). In mouse IMCD3 cells, SEPT2 forms a ring-like structure at the ciliary base. Knockdown of this protein leads to ciliogenesis defects as barrier function is disrupted and more transmembrane proteins enter the ciliary compartment (Fliegauf et al., 2014; Hu and Nelson, 2011; Kim et al., 2010).

In addition to a distinct protein composition at the TZ, phosphoinositide lipids PI (4,5) P2 and PI(4)P are restricted to the proximal ciliary membrane by the ciliary enzyme Inositol Polyphosphate-5-Phosphatase E (Inpp5e). Loss of Inpp5e can cause disruption of Hh signalling pathway and levels of PI(4,5)P2 are increased. Evidence suggests that Inpp5e is involved in both ciliogenesis and ciliary membrane trafficking and that Cep164 protein is required for recruiting this enzyme to the ciliary base (Chávez et al., 2015; Garcia-Gonzalo et al., 2015).

JBTS-MKS-NPHP ciliopathic-associated protein complexes have been identified at TZ/BB, which provided us valuable insights of their roles in cilium biogenesis, maintenance and signalling pathways. However, deeper understanding of molecular mechanisms of transition fibre and TZ function are still required, in addition to identifying components of the TZ and transitional fibres, other ciliary components and their transport machinery. As many of the TZ proteins described above show evolutionary conservation, in my PhD study, I have been using the flagellated

microbial organism; the African trypanosome *Trypanosoma brucei* (a pathogen of significant medical and veterinary importance in Sub-Saharan Africa) to investigate cilium/flagellum formation.

## 1.3 *Trypanosoma brucei* – a model system for studying flagellum biogenesis

### 1.3.1 Background and life cycle

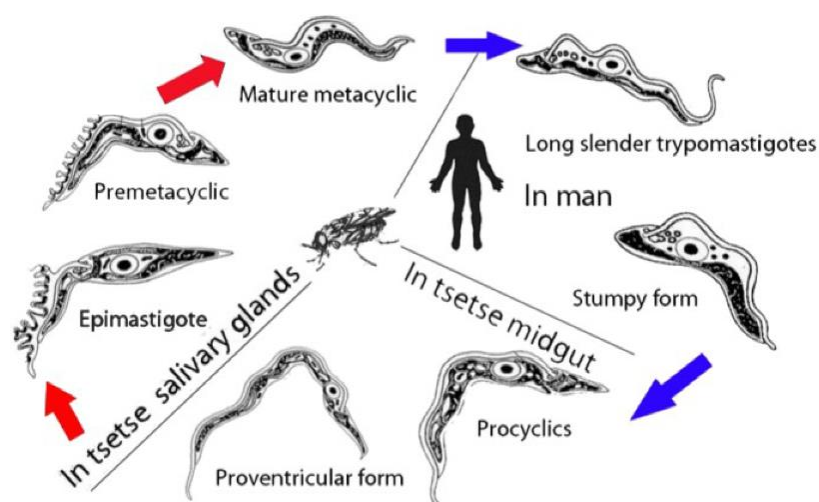
Trypanosomes belong to a group of flagellated protists called kinetoplastids. All members of this group have a unique and distinctive region called the kinetoplast which contains condensed mitochondrial DNA (Matthews, 2005). The transmission of the trypanosome *Trypanosoma brucei*, the causative agent of African trypanosomiasis (also known as African sleeping-sickness), occurs following the bite of an infected tsetse fly (*Glossina* genus). The vector can only be found in tropical and subtropical areas in Africa, though only certain species of tsetse flies transmit disease. Without any effective diagnosis and treatment, the consequences of infection can be fatal. There are two human infective subspecies *T. brucei rhodesiense* and *T. brucei gambiense*. A third subspecies *T. brucei brucei*, causes a livestock wasting disease called Nagana, but does not normally infect humans.

In humans, the transmission of the parasite starts when the tsetse fly takes a blood meal and injects metacyclic trypomastigotes into the human host. Once the parasites enter the host bloodstream, they transform into bloodstream-trypomastigotes (long slender form) and they can survive and move freely to the rest of the body (e.g. central nerve system). Parasites can evade host immune responses due to the expression of the bloodstream stage specific variable surface glycoproteins (VSGs), which are attached to the parasite surface membrane by a glycosylphosphatidylinositol (GPI) anchor (Paulick and Bertozzi, 2008). Long slender form trypomastigotes replicate by binary fission and as the number of parasites in the bloodstream increase the long slender form differentiates into a non-replicative short stumpy form (via quorum-sensing-like mechanism); this form is preadapted for survival within the tsetse fly midgut (MacGregor et al., 2011). This transformation to a cell cycle arrested stage, is required to maximise the possibility of parasite transmission i.e. it ensures the host

does not die before the parasite can be transmitted to another host (Matthews et al., 2004).

Upon transmission to the insect host, short stumpy forms transform into procyclic trypomastigotes in the midgut and re-enter the cell division cycle. From there, procyclic form parasites migrate through peritrophic matrix to the proventriculus, then move to mouthparts, and eventually reach the salivary gland. The short epimastigotes can attach to the epithelium of the salivary gland and undergo replication and differentiation (epimastigotes to metacyclic form), these are released from the gland epithelium into the salivary gland so they are available to infect other hosts (Fig 1.6).

As pathogens of significant medical and veterinary importance, *T. brucei* has been widely studied. However, more recently *T. brucei* has also emerged as an attractive model to study the assembly of the eukaryotic flagellum; due to the conservation of genes/proteins involved in flagellum assembly between human and trypanosomes, and because of the tractable reverse genetic systems available in *T. brucei* (Vincensini et al., 2011).



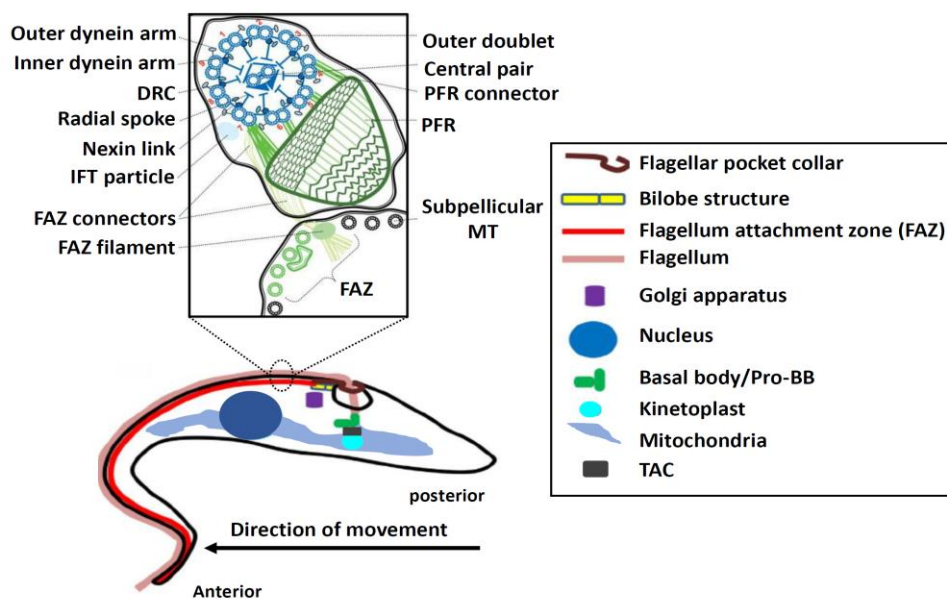
**Fig 1.6 The life cycles of *T. brucei* in tsetse fly and human host.** During their life cycle, the parasites need to go through numbers of developmental and morphological transformation for adapting in different environments. Taken from (Jones et al., 2013) with permission.

### 1.3.2 Cell Architecture

The *T. brucei* cell is vermiform in shape, wider at the cell centre and tapering towards the poles. The anterior and posterior ends of the cell are defined based on the direction of cell movement. The majority of organelles are positioned towards the posterior pole of the cell (Fig 1.7). Trypanosomes assemble a single flagellum, which elongates from the basal body near the posterior end of the cell. The flagellum emerges from the cell body through an invagination known as flagellar pocket (FP) (Field and Carrington, 2009). The flagellum is closely attached to the cell body by a network of electron-dense filaments and membranous connectors known as the flagellum attachment zone (FAZ) but also extends beyond the cell as a short free flagellum (Ralston et al., 2009). As mentioned earlier, all trypanosomatids contain a network of circular DNA inside the kinetoplast within a single elongated mitochondrion (Lukes et al., 2002). The kinetoplast is connected to the mitochondrial membrane by unilateral kinetoplast filaments, and is also attached to the basal body via exclusion zone filaments, together called the tripartite attachment complex (TAC). The TAC is extremely important as it provides and maintains the structural and relative positioning of mitochondrion, kinetoplast, basal body during the cell cycle (Ogbadoyi et al., 2003). As with other eukaryotic cells, *T. brucei* contains a nucleus, Golgi apparatus, lysosomes, endosome, MT cytoskeleton, but also glycosomes that can only be found in Kinetoplastids (which are involved in glycolytic processes) (Michels et al., 2006; Opperdoes and Borst, 1977).

The single *T. brucei* flagellum has a typical 9+2 microtubule-based axoneme, consisting of nine outer doublet microtubules (ODs) surrounding a central pair apparatus (CP) of singlet microtubules. Radial spokes extend from the outer doublet A tubules facing inwards toward to the CP. Both inner and outer dynein arms are also attached to the outer doublet microtubules (Fig 1.7). The paraflagellar rod (PFR) is another unique and characteristic morphological feature in the trypanosome flagellum (and a few related organisms). The PFR is a lattice-like structure that runs alongside of the axoneme, and is attached to axonemal outer doublets 4-7 via fibrous connections. The connection between the PFR and the axoneme only occurs after the flagellum has emerged from

the flagellar pocket and extends to the flagellar tip (Koyfman et al., 2011). In a cross section the PFR is observed to have three domains – proximal, intermediate and distal. The exact function(s) of the PFR are still unclear, but it is known that the PFR is essential and required for normal motility and viability in *T. brucei* (Bastin and Gull, 1999; Bastin et al., 1998). RNAi silencing of two components (PFR1 and PFR2) results in strong reduction in flagellum beating and cell paralysis, but with no obvious defect in axoneme structure (Bastin et al., 1998; Durand-Dubief et al., 2003). Since the first discovery, more than 40 additional proteins have been identified that show association with the PFR through biochemical, bioinformatics and immunological techniques, which provides increasing evidence for a PFR function in regulatory, signaling and metabolic functions (Portman and Gull, 2010).



**Figure 1.7 Cartoon showing the cellular morphology of a procyclic form *T. brucei* cell.** The trypanosome cell in G1 possesses a single large mitochondrion, a kinetoplast containing the mitochondrial DNA that is attached to the basal body via the tripartite attachment complex (TAC). The flagellum extends from the mature basal body exiting the cell via the flagella pocket; the only site of endo/exocytosis in the cell. The flagellum is attached to the cell body via the flagellum attachment zone or FAZ. The cellular DNA is contained in a nucleus. The vermiform shape of the cell is maintained by a highly organised microtubule based cytoskeleton, which underlies the plasma membrane, called the subpellicular corset. The direction of cell movement is indicated, the cell swims with its flagellum beading. The anterior and posterior ends are defined by this directionality of movement. Reproduced from (Ralston et al., 2009; Zhou et al., 2014) with permission.

### 1.3.3 Flagellum biogenesis and cell cycle

In common with other eukaryotes, trypanosomes need to proceed through a defined series of molecular events (coordinated in space and time) to produce two daughter cells. Different morphological forms of *T. brucei* cells can be identified at different stages during the cell division cycle; with observations based on the position of the kinetoplast DNA (kDNA), length of new flagellum growth, and the positioning of the nucleus along the anterior/posterior axis (Fig 1.9). At the beginning of the cell cycle, the trypanosome cell contains a single nucleus, mitochondrion, kinetoplast, basal body and flagellum (Fig 1.8), all of which must be accurately duplicated and segregated prior to cell division. Flagellum assembly is one of the earliest events in the cell cycle in *T. brucei* (Sherwin and Gull, 1989b).

At an early G1 stage, the *T. brucei* cell has a single flagellum with one mature basal body and one pro-basal body (triplet MT  $9_3+0$ ). When the cell enters G1/S phase, the pro-basal body matures and elongates anchors at the flagellar pocket membrane and extends into a transition zone (double MT  $9_2+0$ ), the MTs are connected to flagellar membrane by chalice-shaped filaments to form ciliary necklace (Gilula and Satir, 1972; Vaughan and Gull, 2015) (Fig 1.8). At the end of the TZ, the basal plate is where the elongation and assembly of flagellum (double MT  $9_2+2$ ) starts in *T. brucei*.

The whole process is reliant on IFT-dependent movement of proteins, 19 of the IFT genes have identified as conserved in *T. brucei* (see table 1) (Absalon et al., 2008b; Davidge et al., 2006; Kohl et al., 2003a). IFT proteins are required for flagellum assembly. RNAi knockdown of the retrograde motor HC and IFT88 resulted in cells with short flagella and eventually non-flagellated cells (Kohl et al., 2003a). Later work done by Absalon and co-workers revealed that 14 IFT proteins can be found in the axoneme, as well as at pro- basal body and basal body regions of both old and new flagella. Functional analysis by RNAi knockdown of these IFT proteins showed they are essential for new flagellum assembly but the old flagellum remains unaffected. Furthermore, data supported the idea that these complexes have separable functions in *T. brucei*, as knockdown of IFT complex A or B proteins appeared to have different phenotypes. For instance, when IFT122 and IFT 140 (A complex) were absent, cells still



were able to produce flagella, but these were short, indicating that retrograde IFT transport was defective. In contrast, when complex B proteins (e.g. IFT172 and IFT52) were inhibited, cells completely failed to build an axoneme (Absalon et al., 2008b). Another recent study also revealed that in the *T. brucei* system, when the flagellum is assembled, the function of IFT is to maintain the distribution of flagellum proteins and flagellar beating, but not flagellum length (Fort et al., 2016). In *T. brucei* and all members of the Trypanosomatid family, two distinct genes for IFT dynein heavy chain are encoded in the genome (DHC2.1 and DHC2.2), which form a heterodimer and are essential for retrograde IFT (Adhiambo et al., 2005; Blisnick et al., 2014). DHC2.1 can be found in both anterograde and retrograde IFT, indicating it participates in IFT (Buisson et al., 2013). RNAi knockdown of DHC2.2 leads to defects in flagellum formation (Kohl et al., 2003b). The stability of DHC2.1 and DHC2.2 at the flagellum base depends on the presence of the intermediate dynein chain (DIC5/XBX-1/D1bLIC) and dynein light intermediate chain (DLI1/FAP133/WDR34) (Blisnick et al., 2014). Furthermore, they also showed that in the IFT140 mutant (RNAi) background, IFT dynein components were unable to enter the flagellum. Their data might suggest that IFT dynein motors are assembled in the cytoplasm, migrate to the flagellum base, and the IFT-A complex is required for efficient entry of the IFT dynein in the flagellar compartment, as the data observed in *Chlamydomonas* (Pedersen et al., 2006). On the other hand, IFT-A complex could function as an adapter to allow IFT dynein proteins pass through the TZ (Blisnick et al., 2014).

The axoneme exits the cell body through an invagination of the plasma membrane called the flagellar pocket (FP), which is also the sole site of endocytosis and exocytosis and so is critical for host and parasite interaction and immune evasion (Field and Carrington, 2009). The flagellar pocket collar (FPC) is where axonemal MTs exit the FP, and is a borderline between the FP and pellicular membranes. Moreover, the FPC also serves a structural role, as it tightly connects the flagellar and cell membrane. For example, a calcium-binding, polymer-forming protein named *TbBILBO1* was identified as a component of FPC and is essential for biogenesis of the FPC (Bonhivers et al., 2008). RNAi knockdown of *TbBILBO1* disrupts the formation of the FPC, inhibits the biogenesis of important cytoskeleton structures, induces severe perturbation of the

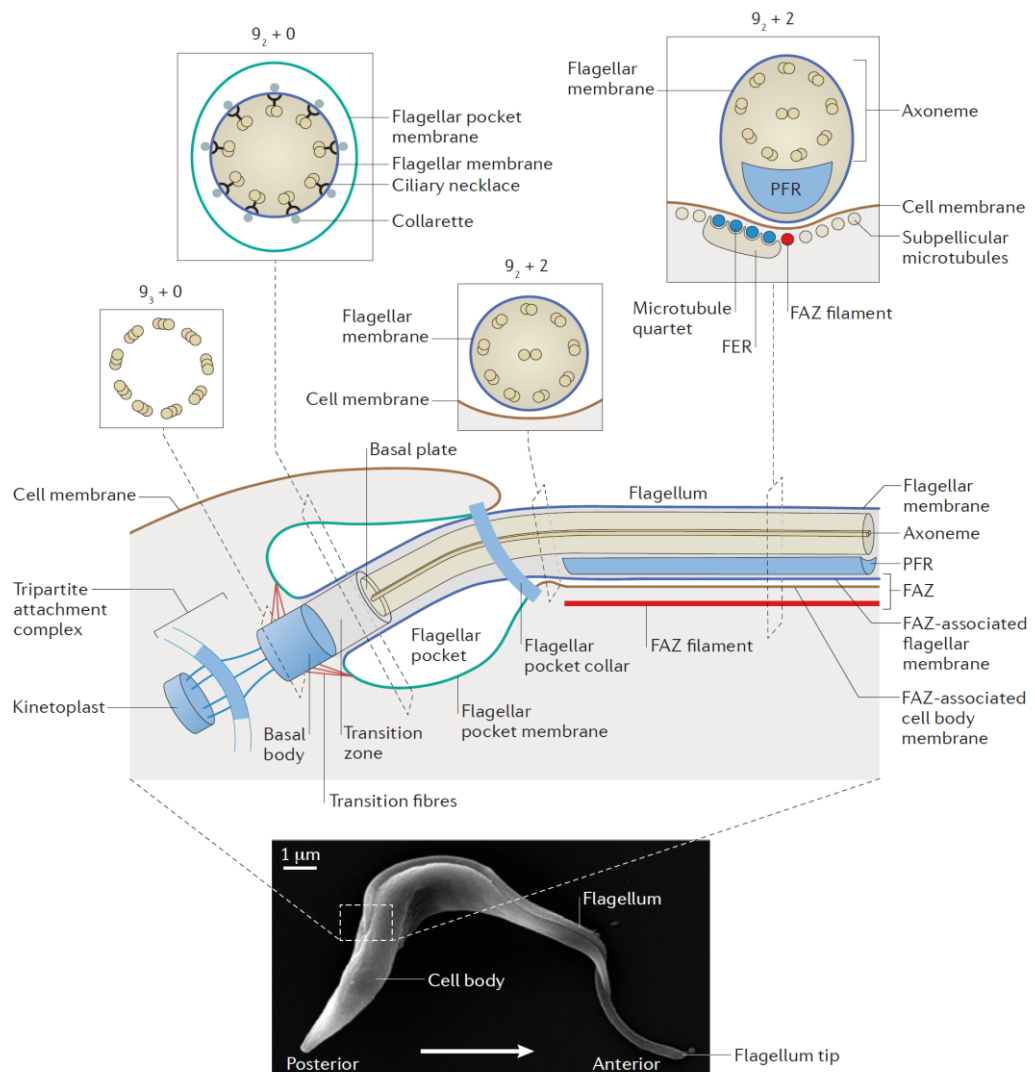
endomembrane system, cell cycle arrest, and is ultimately lethal (Perdomo et al., 2016; Vidilaseris et al., 2015). After this point, the *T. brucei* flagellum extends along the cell body and remains attached until the distal end of the cell body is reached, after which it extends as a free flagellum. The FAZ is a large cytoskeletal structure that connects the flagellum to the cell body after the flagellum exits the FP (Sunter and Gull, 2016). FAZ filaments are located in a gap between sub-pellicular microtubules, which lie beneath the plasma membrane (Vaughan et al., 2008; Zhou et al., 2011). From the posterior end of the cell, the specialised quartet of subpellicular microtubules (MTQ) is on the left of the FAZ filament, and they are associated with smooth ER (Sherwin and Gull, 1989a). The FAZ has been recognised as a key morphogenetic structure regulating both cell length and organelle positioning. There are only a handful of FAZ proteins that have been identified, for example, Flagellum adhesion glycoprotein 1 (FLA1) (LaCount et al., 2002). RNAi knockdown of FLA1 causes flagellar detachment and prevents cytokinesis. Other studies have shown that the depletion of another FAZ protein – the calpain-like protein GM6 (ClpGM6) - produces epimastigote-like morphology cells with long, free flagella and that FAZ length was significantly reduced, with the kinetoplast, basal body, Golgi and FP being repositioned (Hayes et al., 2014). *Trypanosoma brucei* calmodulin (CaM) is localised within the paraflagellar rod (PFR), and is essential for PFR assembly. RNAi knockdown of CaM had no discernible effect on axoneme assembly, but the consequence of PFR assembly defect further results in loss the connection between PFR and axoneme to the FAZ of the cell body. Therefore, similar phenotypes that are observed in FAZ proteins RNAi, can also be seen CaM RNAi. As CaM RNAi mutant reduced the length of FAZ, which cells appeared to be short (Ginger et al., 2013).

As mentioned earlier, *T. brucei* is an attractive model organism to study basal body/TZ proteins as well as flagellum biogenesis and maintenance, because the trypanosome flagellum exhibits the canonical features of the TZ and the trypanosome genome encodes much of the known conserved biology required for TZ function (such as IFT, MKS, and BBS proteins) (Barker et al., 2014; Hodges et al., 2010). A recent study was conducted by Dean and co-workers which identified a large number of TZ and Inversin compartment components. Among all those proteins over one-third of the TZs are

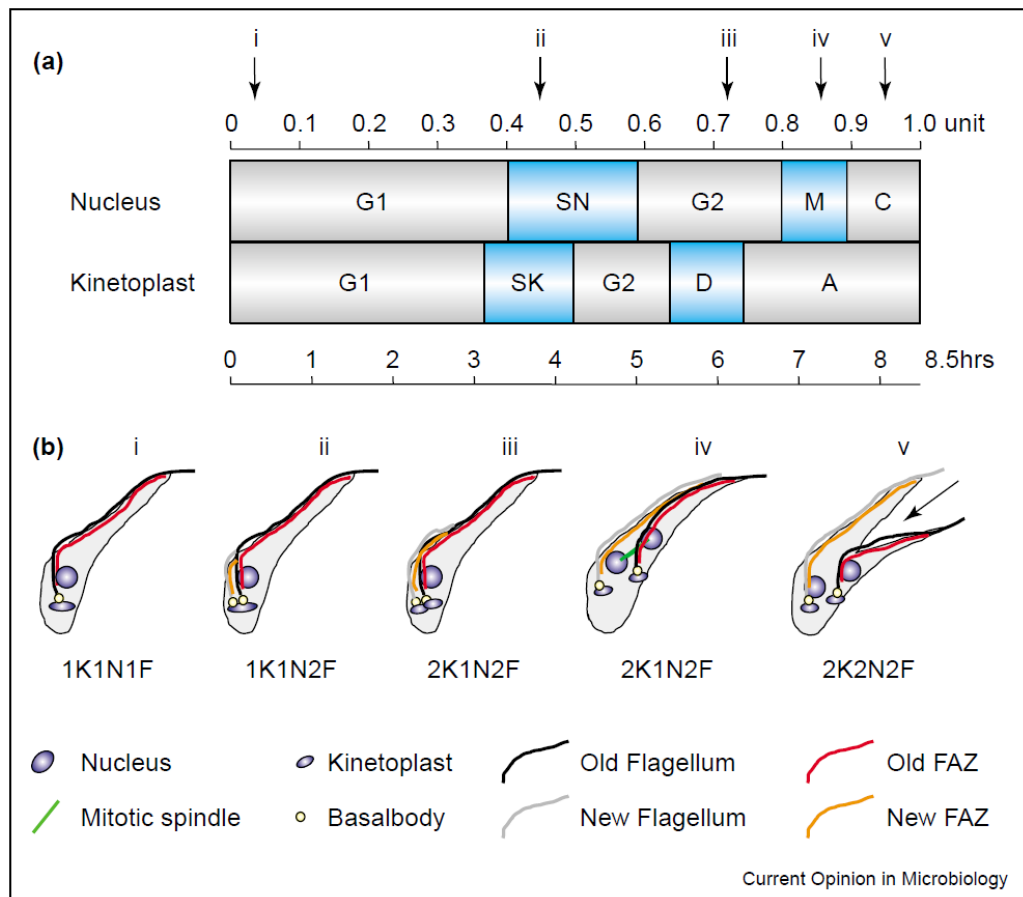
highly conserved across eukaryotes, half of those proteins are kinetoplastid specific. Moreover, about 12 TZs proteins are conserved in *Chlamydomonas* genome. Using fluorescence microscopy and EM also revealed that the BBSome is more distal than the MKS complex within the TZ. Finally, RNAi analysis revealed some of the TZs proteins required in flagellar biogenesis. For example, TZP250 is an orthologue of human ODF3, knockdown of TZP250 leads to severe phenotype, and some are also involved in axonemal structures formation such as MT central pair (e.g. Knockdown of TZP103.8) (Dean et al., 2016).

Like all eukaryotic cells, *T. brucei* has five distinct cell cycle stages - G0, G1, S, G2 and M. However, because *T. brucei* has two genomes (i.e. kinetoplast and nuclear), the cell cycle is more complicated, as both the nuclear and kinetoplast genomes must be replicated and segregated correctly. For the nuclear genome, S-phase is referred as  $S_N$ , and in the kinetoplast  $S_K$ .  $S_K$  starts before  $S_N$  and, as it takes less time to complete, kDNA enters the G2 phase and kinetoplast segregation (D) occurs before the nucleus enters mitosis (Woodward and Gull, 1990). The cell division cycle events take about 8.5 hours in procyclic form cells and each step is highly regulated by molecular regulators of interconnected signal transduction pathways (Fig 1.8-A) (Hammarton, 2007).

At the early G1 phase all *T. brucei* cells have a 1K1N1F configuration i.e. a single nucleus, a single kinetoplast and a single flagellum. As the cell enters the cell cycle the pro-basal body matures to produce a new flagellum. During S phase, the new flagellum extends from the FP and elongates towards the anterior pole of the cell. The distal tip of the new flagellum is physically attached to the old flagellum by an unusual transmembrane mobile junction called the flagellar connector (Briggs et al., 2004b). The kinetoplast starts to divide during G2 phase (2K1N2F) with newly replicated kinetoplast DNA network continuing to move apart during nuclear mitosis. In trypanosomes, cytokinesis is initiated by an invagination of cell membrane between the new and old flagellum at the anterior end of cell. The cleavage furrow progresses unidirectionally along the longitudinal axis of the cell from anterior to posterior pole, until cell abscission is completed forming two daughter cells (Hammarton et al., 2007).



**Fig 1.8. The trypanosome flagellum in procyclic form.** Scanning electron micrograph (EM) of one single procyclic trypanosome (lower image). Schematic diagram shows that the *T. brucei* flagellum, its cross-section and their attached organelles in details. *T. brucei* has conserved canonical “9+2” axoneme structures as well as its unique structures of trypanosomes and closely related organisms. Taken from (Langousis and Hill, 2014) with permission.



**Fig 1.9 Cartoon showing the timing and order of the major morphological events during the cell cycle in procyclic *T. brucei*.** (a) The duration of one cell cycle takes about 8.5 hours. Kinetoplast replication ( $S_K$ ) initiates at ~3 hours, nuclear replication ( $S_N$ ) initiates slightly later and is of longer duration. There is a short period of time when kinetoplast segregation occurs (D). After G2, nuclear mitosis (M) occurs. The “A” phase in kinetoplast cycle is when cell start dividing into two, and basal bodies continue to move apart. (b) The phenotypic changes in trypanosome cells based on the time points shown in (a). The black arrow indicates the direction and position of the cleavage furrow. Diagram taken from (McKean, 2003) with permission, and data based on (Sherwin and Gull, 1989b; Woodward and Gull, 1990)

### **1.3.4 The Importance of the *T. brucei* flagellum**

The contribution of the flagellum to parasite motility and host cell attachment was appreciated for many years; however, functional analyses of flagellar proteins have revealed surprising roles for the flagellum in cell morphogenesis and cell division. In the following sections, I will focus on the importance of flagellum in cell morphogenesis and cell division.

#### **1.3.4.1 Cell morphogenesis and cell division**

Outgrowth of the new flagellum defines the plane of cleavage furrow ingression and its point of initiation (Kohl et al., 2003b; Vaughan, 2010). The flagellum and associated basal body are also critically involved in segregation of the kinetoplast genome. The basal body is connected to the kDNA via the tripartite attachment complex (TAC) via a set of filaments (Ogbadoyi et al., 2003). This physical connection remains intact throughout cell division, during basal biogenesis TAC must go through an extensive re-organisation process. However, precisely how this is orchestrated and regulated is not yet clear (Vaughan and Gull, 2015). During flagellum elongation, the old and new basal body move apart, because the basal body and kDNA are connected by the TAC, which allows the separation of the new and old kDNA and ensures each daughter cell has one basal body and one kinetoplast (Gluezn et al., 2011; Lacomble et al., 2010). The mechanism(s) that drive kinetoplast-basal body movements are unclear. However, it has been proposed that flagellum microtubule elongation is required for the kinetoplast-basal body movement, as the flagellum pushed the kinetoplast-basal body towards the cell posterior, after movement of flagellum tip toward anterior end is stopped (Absalon et al., 2008a; Robinson and Gull, 1991). The new FAZ is nucleated and assembled near the basal body area, before the formation of the new flagellum takes place (Kohl et al., 2003a). The elongation of the new FAZ follows the path of the new flagellum. Cytokinesis is initiated from the anterior tip of the new FAZ filament, proceeds along the longitudinal axis between the old and new flagellum, which appears to be in parallel with the mitotic spindle (Robinson et al., 1995; Vaughan and Gull, 2003). Growing evidence suggests that the new FAZ may provide structural

information for positioning the cleavage plane (Robinson et al., 1995; Sunter et al., 2015; Sunter and Gull, 2016; Vaughan, 2010; Zhou et al., 2015; Zhou et al., 2011).

As mentioned earlier in this section, *T. brucei* has to undergo extensive changes in size and morphology throughout the cell and life cycle. The different forms of *T. brucei* are defined on the relative position of the flagellum to the other organelles, emphasising the importance of flagellum in cell morphology. The length of *T. brucei* varies from 15 to 30 µm long within tsetse flies, and FAZ filament length shows directly correlated to the cell length/size and flagellum length (Kohl et al., 2003a, b; LaCount et al., 2002; Sun et al., 2013). When new flagellum formation is perturbed by RNAi knockdown of *TbDHC1b* or *TbIFT88* short flagella result and the associated FAZ filament also appears to be shorter, which results in cell body size decreased too. Together, the data indicated that new flagellum elongation is required for FAZ assembly (Kohl et al., 2003a).

For several years, the laboratory has been interested in gaining a better understanding of flagellum assembly. First because of the importance that the flagellum has for parasite biology, but secondly because *T. brucei* can be used as a model to study generic (i.e. evolutionary conserved) flagellum assembly processes which could shed light on human ciliopathies. One major area of interest has been the role of a tubulin cofactor C-domain containing (TBCC) protein known as RP2 (retinitis pigmentosa). In the following sections, I provide more detail on the canonical tubulin folding pathway, TBCC domain containing proteins, and specifically the RP2 protein.

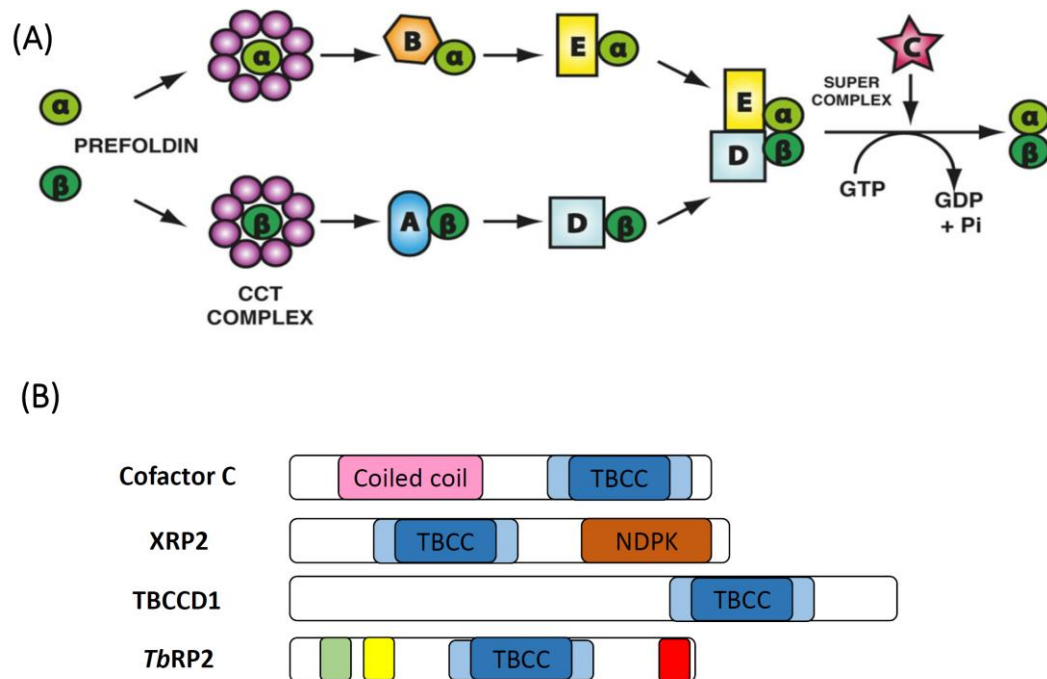
## 1.4 Tubulin folding pathway

The core structural component of eukaryotic cilia and flagella is the 9+2 microtubule-based axoneme constructed from heterodimeric  $\alpha/\beta$ -tubulin. The generation of  $\alpha/\beta$  tubulin heterodimer depends upon a canonical folding pathway, which involves five tubulin-specific cofactors (TBCE-E) to ensure  $\alpha$ - and  $\beta$ -tubulin are correctly folded (See Fig 1.10-A) (Lewis et al., 1997; Tian et al., 1996). At the beginning of the tubulin folding pathway, newly formed  $\alpha/\beta$ -tubulin monomers interact with prefoldin and are transferred to cytosolic chaperonin (CCT) in an ATP dependent manner. ATP hydrolysis by CCT results in conformational changes to produce quasi-native folded tubulin (Llorca et al., 2001; Llorca et al., 1998). The products then bind with tubulin cofactors to carry on the folding process.  $\alpha$ -tubulin specifically interacts with Tubulin Cofactor B (TBCB) and TBCE, whilst  $\beta$ -tubulin independently interacts with TBCE and TBCD. *In vitro*, both TBCE and TBCB are not required for tubulin folding, and it is proposed that their roles are to shift the reaction equilibrium to provide favourable target protein to bind TBCD and TBCE. Based on their position in the pathway, they could act as a reservoir for tubulin folding intermediates. The presence of TBCE and TBCB also ensures *de novo* of tubulin folding is favoured over the back-reaction that involves native  $\alpha/\beta$ -tubulin heterodimer refolding (Tian et al., 1997).

Both  $\alpha$ - and  $\beta$ -tubulin are GTP (guanosine triphosphate)-binding proteins. GTP bound with  $\beta$ -tubulin is freely exchangeable. However, GTP binding with  $\alpha$ -tubulin is non-exchangeable (Spiegelman et al., 1977). GTP bound with TBCE/ $\alpha$ -tubulin and TBCD/ $\beta$ -tubulin complexes help to maintain the stability of the intermediate complexes. Tubulin cofactor C (TBCC) plays a crucial role during the latter part of the tubulin folding pathway. It is proposed that tubulin cofactor C (TBCC) is involved in forming a super-complex with TBCD, TBCE and  $\alpha/\beta$ -tubulin. Biochemical studies suggest that TBCC, TBCD and TBCE behave as GTPase-activating proteins (GAPs) to stimulate  $\beta$ -tubulin (GTP-binding protein) within the super-complex to trigger the hydrolysis of GTP leading to release of the native  $\alpha/\beta$ -tubulin heterodimer components for microtubule incorporation. The hydrolysis of GTP is important as GTP hydrolysis is also



coupled with polymerisation of  $\alpha/\beta$ -tubulin heterodimer and also controls and determines the dynamic behaviour of microtubules (Tian et al., 1999).



**Fig 1.10 (A) Schematic diagram shows the canonical tubulin-folding pathway.** The native  $\alpha$ - and  $\beta$ -tubulin monomers are folded into functional heterodimers by five specific tubulin cofactors A-E. **(B)** TBCC domain containing proteins includes human canonical TBCC, XRP2, TBCCD1 and *Trypanosoma brucei* RP2 (*TbRP2*). The TBCC domain is shown in dark blue, and the light blue indicating two CARP domains (partially overlapped with TBCC domain) usually found in CAP proteins. The coiled-coil regions shown in pink is alpha model, which may be involved in protein–protein interaction, and the nucleotide diphosphate kinase (NDPK) shown in brown found on XRP2 is a phosphocarrier domain. The TOF is in green, and LisH is in yellow and red colour indicates acidic C-terminus. Adapted from (Mori and Toda, 2013).

## 1.5 TBCC domain containing proteins

The majority of eukaryotes encode three tubulin cofactor C (TBCC) domain-containing proteins; canonical TBCC, TBCC-domain containing protein 1 (TBCCD1) and retinitis pigmentosa 2 (RP2). All eukaryotes encode canonical TBCC, which is essential for  $\alpha/\beta$  tubulin folding, but RP2 is found only in ciliated/flagellated eukaryotes, while TBCCD1 is encoded in the genomes of both flagellated and non-flagellated eukaryotes. From a structural viewpoint, TBCCD1 protein is an oddity, as it has been previously reported that a conserved arginine finger residue in the TBCC domain, which has a crucial catalytic function with respect to GAP activity, is missing (Stephan et al., 2007). However, it has been noted in a sequence alignment of *Chlamydomonas* TBCCD1, XRP2 and TBCC orthologues, that an arginine residue appears very close to the functional residue observed in TBCC and RP2 that might stimulate GAP activity (Feldman and Marshall, 2009); but the functionality of this arginine residue has been questioned (Gonçalves et al., 2010). The domain structures of all three human TBCC domain-containing proteins and the *T. brucei* RP2 orthologue (*TbRP2*) are shown at Fig 1.10-B.

### 1.5.1 Canonical TBCC

TBCC consists of three domains; N-terminus domain (alpha module), CARP repeats which occur in many cyclase-associated proteins (CAPs), and a TBCC domain located in the C-terminus of the sequence (Fig 1.10-B) (Garcia-Mayoral et al., 2011). GTP hydrolysis by  $\beta$ -tubulin is essential for biogenesis of native tubulin, and TBCC was first proposed by Tian and colleagues to act as a GTPase-activating protein (Tian et al., 1999). The primary function of TBCC is involved in latter part of the tubulin folding pathway (Fig 1.10-A). Within the supercomplex of  $\alpha$ -tubulin/TBCE and  $\beta$ -tubulin/TBCD, TBCC acts together with TBCD as a GTPase-activating protein, to stimulate the hydrolysis of the exchangeable GTP and allows the native tubulin heterodimer to be released (Tian et al., 1999). The structure of the TBCC domain was

characterised nearly a decade ago (Saito et al., 2007) (PDB: 2YUH), and shown to stimulate GTPase activity of native tubulin in cooperation with TBCD (Bhamidipati et al., 2000). The GAP function of TBCC is proposed to assess 'tubulin quality', thus ensuring that the  $\alpha/\beta$ -tubulin heterodimer is competent to hydrolyse of GTP; essential for microtubule incorporation (Scheffzek et al., 1997). Mutation of the conserved arginine (R262) of canonical TBCC abolishes its GTPase activating protein (GAP) activity suggesting a role in regulation of microtubule polymerization *in vivo* (Bartolini et al., 2002).

The mammalian Arl (ADP-ribosylation factor (ARF)-like) proteins belong to the family of Ras-related small G proteins and share approximately 40 to 60% sequence similarity with ADP-ribosylation factor (Arf) protein (Clark et al., 1993). However, Arf and Arl proteins do not show functional overlap (Amor et al., 2001; Tamkun et al., 1991). Arl2 protein has shown modulate tubulin-GAP activity for both TBCC and TBCD. Arl2 may serve its functions via several ways: (1) down-regulate GAP activity of TBCC, TBCD and TBCE towards  $\beta$ -tubulin *in vitro*. (2) GTPase defective and GTP-binding defective Arl2 mutant showed that the GDP bound form of Arl2 more efficiently interacted with TBCD, which can prevent destruction of tubulin and microtubules due to the TBCD overexpression in cultured cells, but no effect on TBCE (Bhamidipati et al., 2000).

### **1.5.2 TBCCD1**

TBCCD1 (TBCC-domain containing 1) protein is an enigmatic member of the TBCC domain-containing protein family, as it lacks the key catalytic arginine residue responsible for GAP activity (Stephan et al., 2007). To date, only three functional studies have been undertaken on TBCCD1.

The first study was conducted in *C. reinhardtii* model, where the ASQ2 protein encodes the conserved protein TBCCD1 (Feldman and Marshall, 2009). In this study, the majority of cells showed TBCCD1 localises to the region between the centrioles and the nuclear envelope, and in some cells to the centrioles. Mutation in

the ASQ2 gene leads to multiple defects in *C. reinhardtii* cells, such as problems in centriole linkage that cause defects in centriole number and positioning. In addition, cells showed problems in mitotic spindle orientation, but no effect on spindle formation per se (Feldman and Marshall, 2009).

Human TBCCD1 also appears to have multiple locations; centrosome, spindle fibres and basal bodies of both motile and primary cilia (Gonçalves et al., 2010). RNAi knockdown of human TBCCD1 led to cell cycle delay in G1 phase, as well as changing the localisation of the centrosome from the centre of the cell close to the nucleus, to a position about 2µm away from the nucleus. TBCCD1 depletion also caused cilia formation defects and disruption of nucleating microtubules. Microtubules play a critical role in ensuring that the Golgi apparatus (GA) is correctly organised and positioned close to nucleus. Therefore, it came as no surprise that knockout cells showed disorganised and fragmented GA (Rios and Bornens, 2003). The position of the centrosome and GA within a cell is essential for cell polarisation and migration (Vinogradova et al., 2009). The TBCCD1-knockout cells had defects in cell shape, organisation and migration, and the study concluded that TBCCD1 is a centrosomal protein responsible for connecting the centrosome and nucleus (Gonçalves et al., 2010).

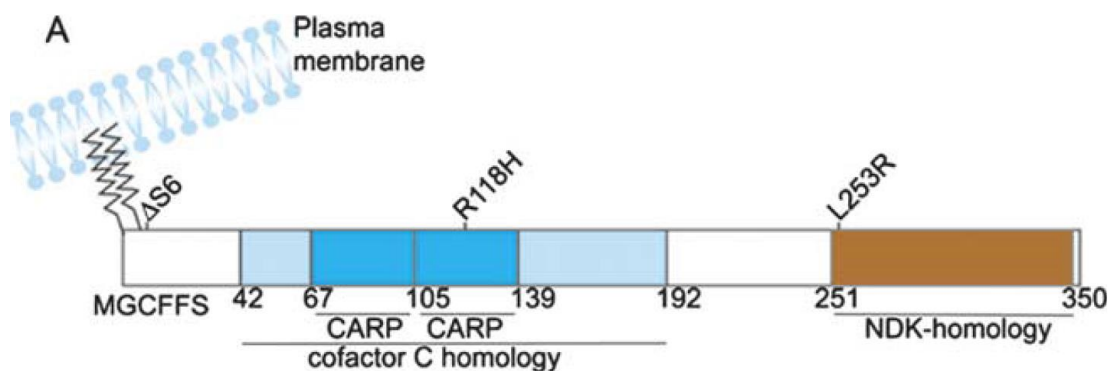
*T. brucei* TBCCD1 has also been investigated, and shown to have an essential role in trypanosome viability (André et al., 2013). The trypanosome orthologue *TbTBCCD1* has three distinct localisations within the cell: the basal body region close to the kinetoplast, the Golgi-associated bi-lobe and anterior end of the cell body. The bi-lobe structure is an enigmatic cytoskeletal feature of *T. brucei*, which localises near the flagellar pocket and was originally suggested to provide a template for biogenesis of the Golgi complex in *T. brucei* (He et al., 2005).

After *TbTBCCD1* RNAi induction the Golgi bi-lobe structure became disorganised and interestingly, the connection between kinetoplast and basal body was affected. Following *TbTBCCD1* ablation *T. brucei* cells were observed with different sized

kinetoplasts, or some cells appeared to have no kinetoplast at all, indicating that loss of *TbTBCCD1* resulting in mis-segregation of the kDNA networks. Based on previously published studies in *Chlamydomonas* and human cell lines (Feldman and Marshall, 2009; Gonçalves et al., 2010) and the observations in *T. brucei* suggested that *TbTBCCD1* had a function in orchestrating the formation of filament-based structures in eukaryotes (André et al., 2013).

### 1.5.3 XRP2

Human retinitis pigmentosa (RP) is a type of inherited retinal dystrophy, which is characterised by progressive degeneration of photoreceptor cells peripheral to the central retina. Patients with RP typically develop night blindness, with vision getting gradually worse until they eventually become completely blind. More than 50 different genetic defects have been identified; autosomal dominant, autosomal recessive, but the most severe form is X-linked RP (XLRP) with mutations in the XRP2 gene responsible for up 15% of all XLRP cases (Schwahn et al., 1998).



**Fig 1.11 Diagram show the domain structure of XRP2 protein.** XRP2 is a 350 amino acid protein; there are two main domains TBCC homology domain and NDK homology domain. R118H, L253R and  $\Delta S6$  are the common mutation take place with RP patients. Taken from (Evans et al., 2006b) with permission.

The human retinitis pigmentosa 2 (RP2) encodes a ubiquitously expressed polypeptide of 350 amino acids and is well conserved across different vertebrate species such as mouse and zebrafish, and loss of the RP2 gene gives similar phenotypes to that seen in humans (Li et al., 2013; Liu et al., 2015). The XRP2 protein contains two domains (Fig 1.11); the N-terminus human XRP2 has dual functional acylation sites (MGCXFSK) and a TBCC domain with a right-handed  $\beta$ -helix folds. Within this TBCC domain region, the tandem repeat cyclase associated proteins (CARPs) motif is also present; the  $\beta$ -helix structure lies within this CARP motif. The N-terminus is post-translationally modified by myristoylation and palmitoylation; modifications that target XRP2 to plasma membranes (Chapple et al., 2000; Schwahn et al., 1998). A disease-causing mutation ( $\Delta$ serine 6) which inhibits dual acylation of XRP2 and prevents localisation to the plasma membrane suggests that membrane localisation is required for RP2 function in photoreceptor cells (Grayson et al., 2002). At the C-terminus of RP2, there is structural homology to nucleoside diphosphate kinases (NDPK) (residues ~251–346) (Chapple et al., 2000; Evans et al., 2006b). However, the NDPK-like domain of XRP2 lacks a critical catalytic residue and isothermal titration calorimetry (ITC) reveals XRP2 does not bind ADP indicating phosphotransferase activity is missing (Kuhnel et al., 2006).

Initial studies suggested that XRP2 function had partial functional overlap with TBCC (Bartolini et al., 2002; Evans et al., 2006b; Grayson et al., 2002; Schwahn et al., 1998; Schwarz et al., 2012c). Bartolini and colleagues have shown that both XRP2 and TBCC simulated GAP activity of native tubulin in concert with TBCD *in vitro*. However, XRP2 was unable to replace canonical TBCC in catalysing tubulin heterodimerisation (Bartolini et al., 2002). XRP2 and TBCC can partially compensate for the loss of *CIN2* (putative homolog of TBCC in *S. cerevisiae*) and act as GTPase activating protein to stimulate the hydrolysis of native tubulin in the presence of TBCD. However, both TBCC and XRP2 fail to replace the function of GAP activity when carrying a mutation affecting the critical arginine finger motif, leading to a defect in the assembly of tubulin heterodimer (Bartolini et al., 2002).

It has been proposed that the arginine finger motif is required for the activity of many GAPs (Scheffzek et al., 1998), and the R118H residue of XRP2 is conserved with R262A residue of TBCC. Intriguingly, mutations in this arginine finger motif completely abolish tubulin-GAP activity in both XRP2 and TBCC. The data indicates that the arginine finger motif plays a key role in tubulin-GAP activity (Bartolini et al., 2002). Mutation of R118H residue in XRP2 is the site of the most common pathogenic missense mutation found in patients with retinitis pigmentosa (Schwahn et al., 1998), which may be due to the loss of GAP activity for XRP2 in the retina (Bartolini et al., 2002).

Grayson and colleagues confirmed the plasma membrane localisation of XRP2 throughout the retina but that XRP2 ablation (Arg120 stop) has no effect on the expression level of TBCC or Arl3. They also demonstrated that Arl3 has a distinct localisation to XRP2, with immunofluorescent labelling showing that Arl3 localised to microtubule structures, including the connecting cilium, suggesting it is a microtubule-associated protein (MAP). The labelling is particularly strong within the rod and cone photoreceptors, suggesting Arl3 may be involved in maintaining the architecture of the photoreceptor. Because Arl3 does not localise to the plasma membrane it is suggested that the interaction between XRP2 and Arl3 is not at plasma membrane (Grayson et al., 2002). The connecting cilium of retina is a specialised microtubule structure, and defects in connecting cilium lead to retinal degeneration. Interestingly, Arl3 can only be found in ciliated and flagellated organisms (Avidor-Reiss et al., 2004). Although, there is no evidence yet to link Arl3 to any human ciliopathies, Arl3<sup>-/-</sup> mice exhibit typical ciliopathic phenotypes includes impaired photoreceptor development but also cysts in kidney, liver and pancreas (die within 3 weeks of birth) (Schrack et al., 2006). In *C. elegans*, Arl3 is also strongly expression in cilium and involved in cilium formation. Knockout Arl3 worms can still form cilia, but when over-expressing the active form of Arl3 lead to defective ciliogenesis, similar result has also been observed in *L. donovani* (Li and Hu, 2011). Those phenotypes are suggesting the role of Arl3 is to act as a negative regulator in ciliogenesis. Intriguingly, depletion of Arl3 can partially rescue ciliogenesis defects in Arl13 mutants by restoring the association between IFT-A and IFT-B. Li and colleagues also demonstrate that Arl3 specifically regulates the

integrity of IFT-B complex and its kinesin motor OMS-3 (KIF17) through the effector deacetylase HDAC6 (Li and Hu, 2011).

Several research groups have solved the crystal structure to identify the function of XRP2 and Arl3, and their binding regions (Kühnel et al., 2006; Veltel et al., 2008). The data has determined that XPR2 is a negative regulator of Arl3 and that the N-terminus and  $\beta$  helix domain of XRP2 are essential for XRP2-Arl3 interaction. Based on affinity measurements, Glu138 and Arg118 are two residues located on the surface of the  $\beta$  helix, mutation in either Glu138 or Arg118 residues lead to XRP2 reduce affinity toward Arl3 drastically. The results suggest that failure of XRP2 to bind Arl3 may cause human retinitis pigmentosa (Kühnel et al., 2006). Further studies by Veltel and colleagues, presented a crystal structure of the Arl3-GppNHp-RP2 complex. Biochemical analysis of this complex represents a 90,000-fold stimulation of the GTPase reaction, indicating that XRP2 is an efficient GAP for Arl3, with structural features similar to other GAPs (Veltel et al., 2008).

Human Arl3 and Arl2 share significant homology. In 2006, Zhou et al., revealed that both Arl2 and Arl3 are localised at the centrosome, and expression of mutant Arl2 leads to defects in microtubule formation and cell cycle arrest in M phase, which may indicate functions in regulating tubulin polymerization at centrosomes. However, no obvious phenotypic morphology was observed following Arl2 knockdown. Data suggested that Arl2 functions as a regulator of tubulin folding and microtubule integrity, but this function may not be essential (Zhou et al., 2006). The function of Arl2 also been studied using human adenocarcinoma cell lines and shown that changing Arl2 protein levels results in significant changes to tubulin content in polymerisable soluble heterodimers. The results suggested that Arl2 proteins are associated with microtubule dynamic instability (Beghin et al., 2007).



Several interactors/effectors of Arl3 that have been identified. XRP2 act as a GAP for Arl3 and involved in vesicle transport and docking to the base of the human photoreceptor cells (Evans et al., 2010). Other effectors, including PDE6 $\delta$  and UNC119, share a similar hydrophobic lipid binding pocket, which can bind with GTP-bound Arl3, and together with XRP2 regulate the delivery of lipid-modified cargo to the cilium (Schwarz et al., 2012a). In addition to a role in cilia formation, Arl3 proteins also function in ciliary signalling pathways via key signalling receptors/molecules; for example in *C. elegans*, the ciliary proteins Polycystin-1 and Polycystin-2, and in Arl3 knockout worms, polycystins fail to target to cilia (Barr et al., 2001; Barr and Sternberg, 1999). In mammalian model, RNAi ablation also disrupts the transportation of Gli3 that participate in Hedgehog pathway (Lai et al., 2011).

Some evidence suggests that Arl3-GTP will bind the effector protein Human retinal gene 4 (*HRG4*, also called UNC119; uncoordinated 119 protein) to form a temporary ternary complex Arl3:UNC119:XRP2 before XRP2 stimulates the hydrolysis of Arl3-GTP (Veltel et al., 2008). This ternary complex traffics cargo to the membrane (Ismail et al., 2011; Wright et al., 2011). Mori and Toda were able to identify Tbc1 and Alp42 (orthologs of human TBCC and Arl2 respectively) in fission yeast model system and demonstrated that Tbc1 is not only act as GAP for the small GTPase Alp42, but also involved in heterodimerisation of tubulin (Mori and Toda, 2013).

Apart from plasma membrane, XRP2 also localises at the basal body, ciliary apparatus of photoreceptor cells, Golgi, and periciliary ridge (Evans et al., 2010; Hurd et al., 2010; Hurd et al., 2011). The function of those organelles are closely related to control of protein export from the outer segment of photoreceptors; which suggests XRP2 may play role in vesicle trafficking and IFT (Maerker et al., 2008; Roepman and Wolfrum, 2007; Schwarz et al., 2017). In mammal, depletion of XRP2 causes dysregulation of Arl3 and disturbs Golgi complex cohesion. Loss of Arl3 and Kif3a (IFT anterograde motor) also leads to dispersal of Golgi complex and morphology changes. Depletion of XRP2 and dysregulation of Arl3 disturbs vesicle cycling cargo from the Golgi

complex to the cilium, and dispersal of IFT20 from the peri-basal body pool. Therefore, results suggest GAP activity of XRP2, and its regulation of Arl3, is crucial for maintaining stability of Golgi complex cohesion and vesicle trafficking in photoreceptor pericentriolar cells (Evans et al., 2010; Follit et al., 2010).

Recent work by Schwarz and co-workers has demonstrated that Kif17 and Kif7 (IFT anterograde motors) are novel interacting partners of Arl3 and its GAP-RP2, but that these two kinesins do not interact. The localisation of Kif17 and Kif7 to the cilia tips depends on the presence of Arl3 and RP2, as RNAi knockdown of either RP2 or Arl3 significantly reduced the Kif17 and Kif7 level at the cilia tips, without affecting the other localisations of Kif17 such as nuclear or cytoplasm (Schwarz et al., 2017). Kif7 is a kinesin family-4 member protein, which is not thought to be essential for IFT or the trafficking of Hh pathway proteins into cilia, but rather regulates the dynamics of microtubule plus ends and thereby controls the cilium architecture and organisation of primary cilia tips (He et al., 2014). Loss of either Kif7 or Kif17 leads to the levels of tubulin polyglutamylation significantly reduced, but total cellular polyglutamylation levels remain unaffected, suggesting that Kif7 and Kif17 specifically stabilise cilia microtubules. Interestingly, RNAi knockdown of Arl3 effector proteins, such as UNC119 and PDE6D did not affect Kif17 and Kif7 at the cilia tips. Together their data suggests that that RP2 and Arl3 regulate the trafficking of specific kinesins to cilia tips, but not dependent upon lipidated protein trafficking (Schwarz et al., 2017).

Apart from well-accepted XRP2-interacting partner protein Arl3, XRP2 also interacts with other proteins such as N-ethylmaleimide sensitive factor (NSF) (Holopainen et al., 2010). Proteomic and biochemical analysis reveals that this interaction is mediated by the N-terminal domain of NSF, and can be observed in both retinal and cultured embryonic kidney (HEK293). Mutations in XRP2 ( $\Delta$ I137 and E138G) lead to this interaction failing. Immunofluorescence analysis shows that XRP2 and NSF are co-localised in photoreceptors and other retinal cells, with a more intense signal within the synaptic region, as well as inner and outer segments beneath the connecting

cilium. This might indicate that the function of XRP2 is not only as a regulator of Arl3, but also interacts with NSF to traffic NSF to the cilium and synaptic region of photoreceptors (Holopainen et al., 2010).

To date there are three main possible functions of XRP2; (1) Tubulin quality control (Bartolini et al., 2002); (2) vesicle trafficking (Schwarz et al., 2012b; Schwarz et al., 2017; Schwarz et al., 2012c); (3) The demonstration that of XRP2 localised to the nucleus when HeLa or ARPE-19 cells were treated with DNA damaging agents (Yoon et al., 2006) seems to be at odds with all other published work on XRP2.

#### **1.5.4 *TbRP2***

The *T. brucei* RP2 ortholog was identified by phylogenetic analysis (Figure 1.11-B), and localised to the transitional fibres of the mature basal body (BB) where it is involved in flagellum assembly (Stephan et al., 2007). Initially, the function of *TbRP2* was proposed to be in ensuring the 'quality' of tubulin imported into the cilium/flagellum (i.e. as GAP for  $\beta$ -tubulin) as *TbRP2* RNAi-ablation caused axonemal microtubule defects, as well as a loss of YL1/2 signal at the BB. YL1/2 is a monoclonal antibody classically used to detect tyrosinated  $\alpha$ -tubulin, and so the loss of basal body YL1/2 labelling observed in *T. brucei* was taken as evidence for RP2 depletion affecting the recruitment of unpolymerised tubulin to the basal body.

However, more recently the function of *TbRP2* has been re-evaluated (Andre et al., 2014). In this study, the targeting of *TbRP2* to the basal body was investigated by expressing a series of *TbRP2*::YFP fusion proteins that contained; (1) just the N-terminus of *TbRP2* (up to but not including the TBCC domain), (2) the N-terminus of *TbRP2* and the TBCC domain, and (3) almost full length *TbRP2* protein that lacked just the acidic C-terminus. Bioinformatic analysis revealed that the N-terminal region encodes twinned TOF- LisH motifs and construct-1 (i.e. just the N-terminus of *TbRP2* containing the TOF-LisH motifs) confirmed that the first 133 amino acids at N-terminus

of *TbRP2* were required for basal body targeting and retention at the transitional fibres. Subsequent site directed mutagenesis experiments confirmed the importance of the TOF-LisH motifs. Interestingly, the use of the twinned TOF-LisH motifs to target RP2 to the basal body appears to be trypanosomatids specific (although TOF-LisH motifs are conserved in other eukaryotic conserved basal body/centrosome located proteins).

In order to determine the GAP function of *TbRP2*, constructs were also made to express myc-tagged *TbRP2*<sup>R248H</sup> mutant protein (*TbRP2*<sup>R248H</sup>::myc). The R248H mutation affects the conserved arginine finger within the TBCC domain required for GAP activity in canonical TBCC and XRP2 (Ding et al., 2008). Compared with the control cell line (*TbRP2*::myc) cells expressing *TbRP2*<sup>R248H</sup>::myc had shortened flagella indicating that over-expression of *TbRP2*<sup>R248H</sup>::myc (i.e. protein lacking GAP function) had a dominant negative effect on flagellum assembly.

To investigate potential functional overlap between XRP2 and *TbRP2*, the human XRP2 protein was targeted to the basal body by creating a chimeric construct that paired XRP2 coding sequence with the sequence encoding the twinned LisH-TOF motifs present at the N-terminus of the *TbRP2* protein. After *TbRP2* RNAi induction, flagellum length was reduced, indicating that XPR2 targeted to the basal body could not functionally compensate for loss of *TbRP2* and restore flagellum length (Andre et al., 2014).

Interestingly, loss of *TbRP2* was also shown to affect the recruitment of both *TbMKS1* and *TbMKS6* to the TZ, suggesting that the flagellum assembly defect resulting from loss of *TbRP2* may also reflect problems in MKS complex assembly. Together, the available data suggested *TbRP2* could have multiple roles in flagellum formation e.g. tubulin processing, general protein trafficking for flagellar assembly, as well influencing TZ protein complex assembly pathways such as those involved in the recruitment of IFT machinery (Andre et al., 2013; Garcia-Mayoral et al., 2011; Schwarz et al., 2012c).

## 1.6 Project aims

Work to date indicates that *TbRP2* is localised to the mature basal body, by virtue of paired N-terminally located TOF-LisH motifs, and is essential for flagellum biogenesis. Homology to XRP2, the presence of a canonical TBCC domain, and the dominant negative phenotype that results from over-expression of mutant (i.e. GAP-negative) protein, strongly supports the view that *TbRP2* is a bona fide GAP protein.

In this project, I set out to identify and investigate proteins that interacted with *TbRP2* or were in close proximity to the protein. To achieve this *TbRP2* was expressed as a fusion with bacterial biotin ligase and by using streptavidin affinity purification biotinylated proteins were purified identification by mass spectrometry; an approach known as BioID. It was envisaged that this approach would identify *TbRP2*-interacting proteins that might provide insight into (i) the molecular client of *TbRP2* GAP activity; (ii) basal body and/or transitional fibre proteins to which *TbRP2* is tethered (in a detergent resistant manner); and (iii) proteins responsible for targeting (in a TOF-LisH dependent fashion) to the mature basal body. The results of this BioID interrogation and subsequent validation of putative *TbRP2* proximal proteins is presented in this thesis.

# Chapter 2 Material and Methods

## 2.1 Chemicals and Reagents

### Melford Laboratories

Agarose, dithiothreitol (DTT), ethylenediaminetetraacetic acid (EDTA), glycerol, glycine, piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES), phleomycin, Phosphate-buffered saline (PBS), sodium chloride, sodium dodecyl sulphate (SDS), PMSF, Benzamidine.

### Sigma-Aldrich

Ampicillin, acetic acid, acrylamide, acetonitrile, ammonium bicarbonate (ABC), bovine serum albumin (BSA), doxycycline, disodium phosphate, hemin, hydrochloric acid, imidazole, iodoacetamide, magnesium acetate, monopotassium phosphate, paraformaldehyde (PFA), sodium carbonate, sodium citrate, sodium hydroxide, tween 20, Triton-X, Urea, tetramethylethylenediamine (TEMED), ammonium persulphate, Tris-acetate, Trypsin Singles: Proteomics Grade, Enzyme, Blastcidin, Doxycycline, Hygromycin, Kanamycin, Phleomycin, Puromycin

### Fisher Scientific

Calcium acetate, ethanol, methanol, magnesium chloride, magnesium sulphate,

### Thermo Scientific

Miniprep Kit, PCR Purification Kit, Gel Extraction Kit, Elution buffer, Halt Protease Inhibitor single-use Cocktail, Reddymix®PCR Master Mix, Streptavidin Magnetic Beads (Pierce™), Formic Acid, Trifluoroacetic acid

### Duchefa Biochemie

LB broth, LB agar, potassium chloride

### BioSera

Foetal bovine serum (FBS)

**Biorad**

Precision Plus (all blue) Pre-stained Protein Standards

**Fluka Biochemika**

Nonidet P-40 (NP-40)

**Fermentas**

Restriction enzymes and their buffers

**GIBCO**

SMD-79 powder

**Formedium**

Glucose

**Vector Laboratories**

Vectashield with 4, 6 diamidino-2-phenylindole (DAPI)

**Promega**

T4 DNA Ligase and buffer, pGEMT-Easy vector system

**Qiagen**

High fidelity PCR mix

**Marvel**

Skimmed milk powder

**Acros Organics**

Orange G

**Roche**

Anti-GFP antibody

## 2.2 Solutions and Buffers

- **Agarose gel:**
- **Blocking buffer for Immunofluorescence:** 1xPBS, 0.05% Tween-20, 1% Bovine serum albumin
- **Blocking buffer for Immunoblotting:** 1xPBS, 0.05% Tween-20, 5% or 1% skimmed milk powder
- **Dithiothreitol (DTT):** 1M stock solution
- **DNA loading buffer (x6):** 0.15% Orange G, 30% glycerol in Milli-Q
- **Elution buffer:** 10mM tris (hydroxymethyl) amino methane (Tris-HCl), pH8.5.
- **Ethanol:** absolute
- **Hemin Stock\* (for SDM-79):** 4% Hemin in 0.1 M NaOH
- **Immunofluorescence slide fixing solution:** 3.7% (w/v) Paraformaldehyde in 1xPBS
- **Insoluble Binding Buffer:** 8M Urea, Tris-HCL (pH8.0), 500mM sodium chloride, 0.02% Triton-X, 20mM Imidazole, 10% glycerol
- **Insoluble Elution buffer:** 8M Urea, Tris-HCL (pH8.0), 500mM sodium chloride, 0.02% Triton-X, 500mM Imidazole, 10% glycerol
- **Inhibitor Cocktail:** Halt Protease Inhibitor single-use Cocktail, 200 mM TLCK, 200 mM PMSF, 1 M Benzamidine
- **LB Agar:** 30g LB agar powder made up to 1L with Milli-Q
- **LB Medium:** 25g LB Broth powder made up to 1L with Milli-Q
- **Methanol:** absolute
- **PBS:** 1 tablet per 200ml Milli-Q water (137mM NaCl, 2.7mM KCl, 10mM Na<sub>2</sub>HPO<sub>4</sub>, 2mM KH<sub>2</sub>PO<sub>4</sub>)
- **Paraformaldehyde Fix for Slides:** 3.7% (w/v) PFA in 1x PBS
- **PEME 1% NP40:** 1% nonidet-P40 in 0.1M piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES), 2mM ethylene glycol tetraacetic acid (EGTA), 1mM magnesium sulfate, 0.1M ethylenediaminetetraacetic acid (EDTA), pH6.9
- **Procyclic trypanosome media:** SMD-79 powder (Brun and Schönenberger, 1979) dissolved in 5L of Milli-Q with 2 g/L sodium bicarbonate (pH 7.3), and



sterile filtered 450 mL into each autoclaved bottle. 7.5 mg/L haemin (10 mg/ml stock in 0.1 M NaOH) and 50 ml heat inactivated foetal calf serum (filtered) were added into the bottle before use. Stored the complete SMD-79 media (500 mL) at 4°C

- **Reddymix®PCR Master Mix:** 0.625 units ThermoPrime Taq DNA Polymerase, 75mM Tris-HCl (pH 8.8 at 25°C), 20mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1.5mM MgCl<sub>2</sub>, 0.01% (v/v) Tween-20, 0.2mM each of dATP, dCTP, dGTP and dTTP
- **SDS loading buffer:** 100mM Tris-HCL pH 6.8, 4% SDS (electrophoresis grade), 0.20% Bromophenol Blue, 20% glycerol, 200mM dithiothreitol solution (DTT)
- **SDS-PAGE 10% resolving gel:** 30% Acrylamide mix, 1.5M Tris-HCl (pH8.8), 10% SDS, 10% ammonium persulphate, 0.1% N,N,N',N',-tetramethylethylenediamine (TEMED)
- **SDS-PAGE 5% stacking gel:** 30% Acrylamide mix, 1M Tris-HCl (pH6.8), 10% SDS, 10% Ammonium persulphate, 0.1% N,N,N',N',-tetramethylethylenediamine (TEMED)
- **SDS-PAGE running buffer (5x):** 125mM Tris, 1.25M glycine, 0.5% SDS
- **Sodium acetate solution:** 3M pH5.2
- **S.O.C medium:** 2% Tryptone, 0.5% Yeast Extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>, 20 mM Glucose
- **TAE (Tris-Acetic acid-EDTA) buffer:** 40mM Tris-acetate, 1mM EDTA, pH8
- **Transfer buffer for Immunoblot (1x):** 48mM Tris-HCl, 39mM glycine, 0.037% SDS, 20% methanol
- **Wash buffer:** 1xPBS with 0.05% Tween-20
- **ZMG buffer:** 132mM Sodium Chloride, 8mM Potassium chloride, 8mM Disodium phosphate, 1.5mM monopotassium phosphate, 0.775mM Magnesium acetate, 0.063mM Calcium acetate (pH7.5)

## 2.1.1 Solutions and Buffers for Mass Spectrometry (BioID/SILAC) (OB-on beads digest, IG-in gel digest)

- **IG-ABC Solution:** 100 mM Ammonium Bicarbonate
- **IG-DTT/ABC Solution:** 10 mM DTT in 100 mM Ammonium Bicarbonate
- **IG-IAA/ABC Solution:** 50 mM Iodoacetamide in 100 mM Ammonium Bicarbonate
- **IG-Extraction Buffer:** 1:2 (v/v) 5% formic acid: 100% Acetonitrile
- **IG-ABC/MeCH Solution:** 10 mM ammonium bicarbonate contains 10% (v/v) acetonitrile
- **OB-ABC Solution:** 50 mM ammonium bicarbonate
- **OB-Elute Solution:** 0.1% Trifluoroacetic acid, 70% Acetonitrile
- **OB-IAA Solution:** 8 M Urea, 1% SDS, 50 mM Iodoacetamide
- **OB-Solubilisation Buffer:** 8 M Urea, 1% SDS
- **OB-TFA Solution:** 0.1% Trifluoroacetic acid
- **OB-TFA/MeCN Solution:** 0.1% Trifluoroacetic acid, 70% Acetonitrile
- **Trypsin Singles:** Proteomics Grade, Enzyme
- **Procyclic trypanosome media:** SMD-79-SILAC

Ingredient	Product	500 mL
SDM-79-SILAC	Cassion Labs custom synthesis	
Heamin stock, 10 mg/mL in 0.1M NaOH	Sigma H9039, 4 °C	375 µL
Sodium bicarbonate	Sigma S5761, RT	1 g
D-glucose	100 x, 9.3 g / 50mL	5 mL
D-glucosamine	100 x, 0.25 g / 50mL	5 mL
L-hydroxyproline	100 x, 10 mg / 50mL	5 mL
L-proline	100 x, 3.08 g / 50mL	5 mL
i-inositol	100 x, 6.5 mg / 50mL	5 mL
L-methionine	100 x, 0.45 g / 50mL	5 mL
L-serine	100 x, 0.36 g / 50mL	5 mL

Adjust to pH 7.3 with NaOH, sterile filter		
Glutamax 1, 100x	Gibco 35050038, -20 °C	75 mL
Penicillin-Streptomycin, 100x	Gibco 15140-122, -20 °C	5 mL
Foetal bovine serum, heat inactivated, FS (add in the hood)	Sigma F4135, -80 °C	5 mL

**Table 2.1 Recipe for making SDM-79-SILAC media for BioID experiments.** Complete SILAC media was filter sterilised the microbiological safety cabinet.

- SILAC labels:**

Label (100mg/ml, -20 °C)	Mw	30% SDM-79
L-Arg.HCl (R0)	210.6	64.5 mg/L
L-Arg.HCl (R6- <sup>13</sup> C <sub>6</sub> )	216.6	66.3 mg/L
L-Arg.HCl (R10- <sup>13</sup> C <sub>6</sub> , <sup>15</sup> N <sub>4</sub> )	220.6	67.5 mg/L
L-Lys.HCl (K0)	182.6	21.6 mg/L
L-Lys.2HCl (K4- <sup>2</sup> H <sub>4</sub> )	223.1	26.7 mg/L
L-Lys.2HCl (K8- <sup>13</sup> C <sub>4</sub> , <sup>15</sup> N <sub>4</sub> )	227.1	27.3 mg/L

**Table 2.2 Recipe for making light-, medium- and heavy-labelled amino acids,** which were added to the SDM-79-SILAC media for SILAC labelling experiments. Arg – Arginine; Lys – Lysine

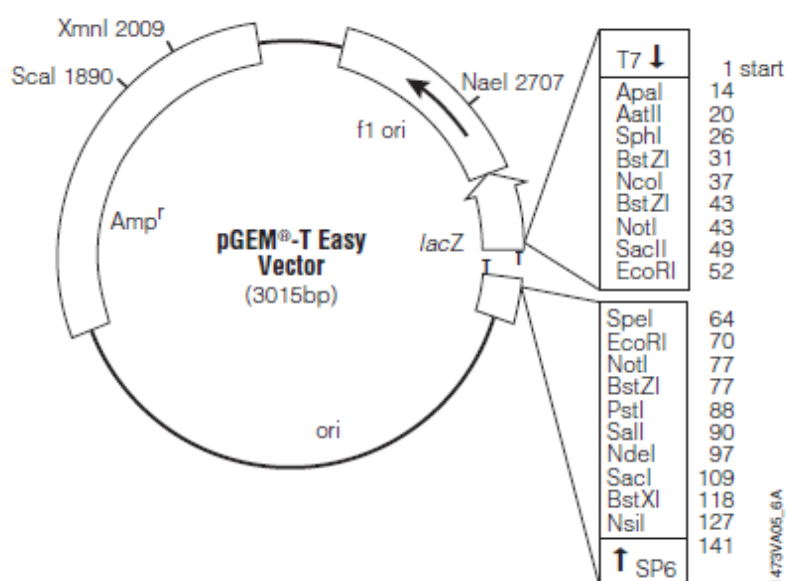
## 2.3 Antibiotics and drug stock solutions

Antibiotic	Stock Concentration	Working Concentration	Use
<b>Ampicillin</b>	100 mg/mL	100 µg/mL	Selective marker for <i>E. coli</i>
<b>Blasticidin</b>	10 µg/µL	10 µg/mL	Selection for <i>T. brucei</i> cell lines

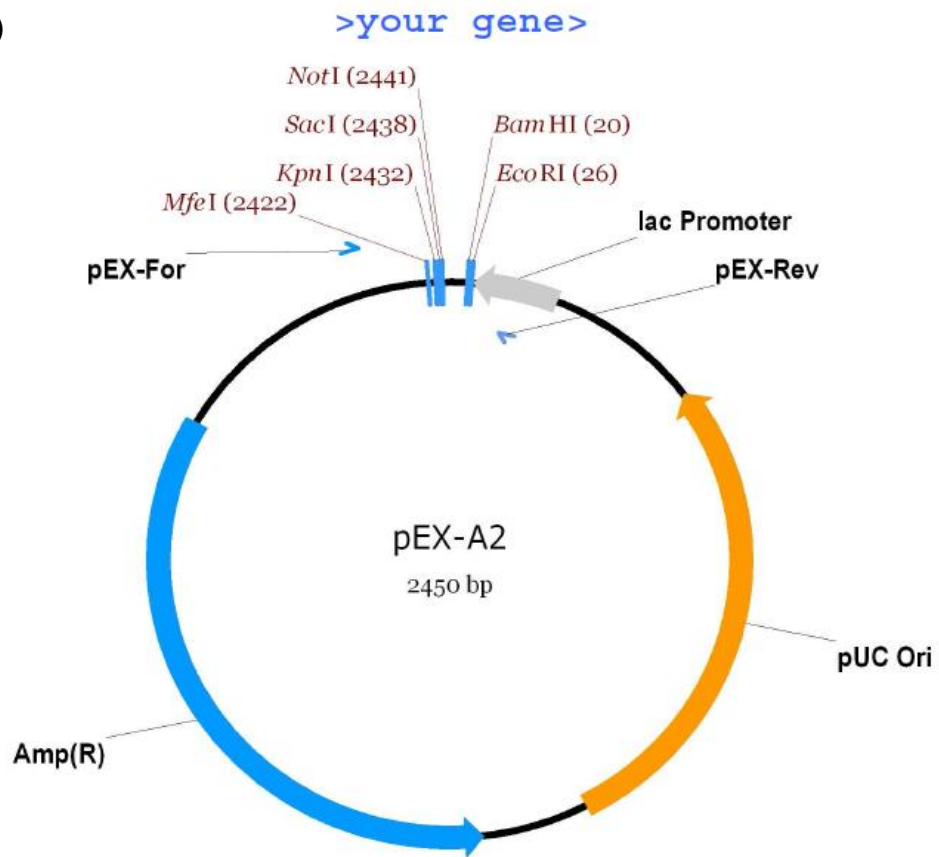
<b>Doxycycline</b>	1 mg/mL	1 µg/ml	Inducible expression for <i>T. brucei</i>
<b>Hygromycin</b>	20 µg/µL	50 µg/mL	Selection for <i>T. brucei</i> cell lines
<b>Kanamycin</b>		10 µg/ml	Selective marker for <i>E. coli</i>
<b>Phleomycin</b>	7.5 mg/mL	3 µg/mL	Selection for <i>T. brucei</i> cell lines
<b>Puromycin</b>	7.5 mg/mL	2 µg/mL	Selection for <i>T. brucei</i> cell lines

**Table 2.3 Antibiotics and drugs used during this study.** All solutions involved in *T. brucei* cell culture was filter sterilised before use.

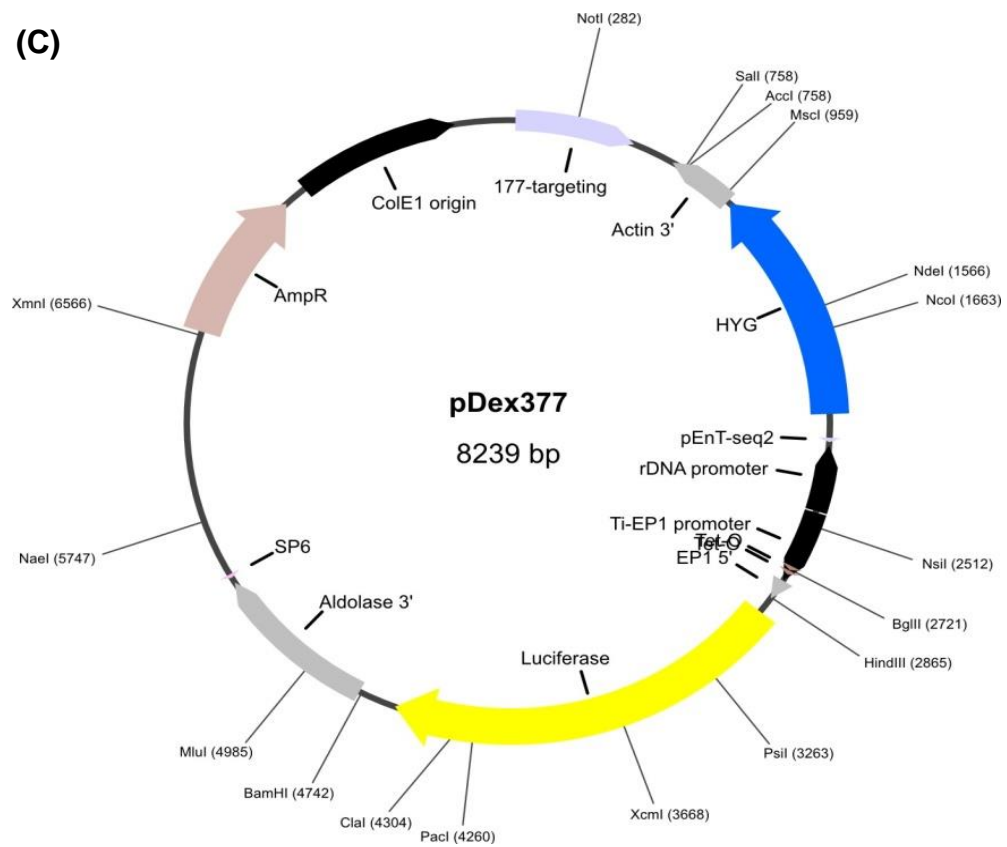
## 2.4 Vector maps

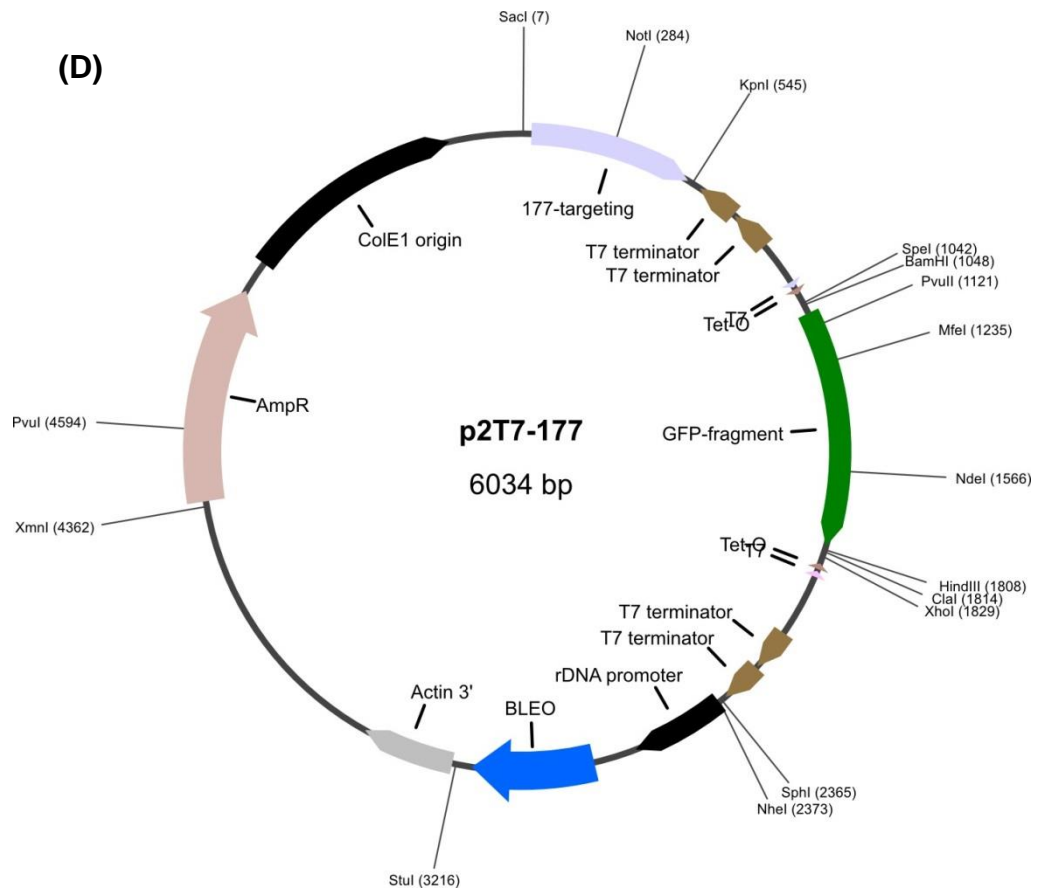


(B)



(C)





**Figure 2.1 Vector maps.** (A) pGEM®-T Easy vector (Dawson and House, 2010). (B) The *TbRP2*<sup>Δ134-463</sup>::myc::BirA\* was generated by gene synthesis (MWG Eurofins) and inserted into the pEX-K4 vector. (C) pDex-377 vector used for constitutive expression of myc-tagged proteins in procyclic trypanosomes (Kelly et al., 2007). (D) p2T7-177 vector was used for a doxycycline-inducible expression of RNAi (Wickstead et al., 2002).

## 2.5 Antibodies and PCR primers

**Table 2.5 Antibodies used for immunofluorescence microscopy**

Primary antibody	Epitope	Raised in	Dilution	Secondary antibody	Dilution	Company
Anti-Myc	Myc	Mouse	1:200	Anti-mouse IgG rhodamine conjugate	1:200	Millipore
L8C4	Paraflagellar Rod	Mouse	1:10	Anti-mouse IgG FarRed	1:200	Invitrogen
Anti-TbRP2	TbRP2	Rabbit	1:50	Anti-rabbit IgG FITC conjugate	1:40	Sigma

**Table 2.6 Antibodies used for Immunoblotting detection**

Primary antibody	Epitope	Raised in	Dilution	Secondary antibody	Dilution	Company
Anti-Myc	Myc	Mouse	1:1000	Anti-mouse IgG HRP conjugate	1:80 000	Novagen
KMX-1	beta-tubulin	Mouse	1:1000	Anti-mouse IgG HRP conjugate	1:80 000	Novagen
Anti-GFP	YFP-tag	Mouse	1:1000	Anti-mouse IgG HRP conjugate	1:80 000	Roche

**Table 2.7 Streptavidin conjugated protein used for IF and IB**

	<b>Dilution</b>	<b>Company</b>
Streptavidin conjugated to TRITC	1:200	Roche
Streptavidin conjugated to HRP	1:10,000	Abcam

## **2.6 Oligonucleotide primers**

Oligonucleotide primers were ordered from Eurofins MWG Operon. Forward and reverse primers for pPot tagging were designed using the pPOT design programme (Dean et al., 2015), and RNAi primers designed using RNAit (Redmond et al., 2003). All primers were re-suspended in filtered Elution Buffer (Thermo Scientific) at a concentration of 100µM (stock) and stored at -20°C. Oligonucleotides were diluted to 10 µM prior to use in PCR.



**Table 2.8 pPOT primer sequences used in PCR reactions**

Primer ID	Direction	Nucleotide sequence
<i>Tb</i> ARL3 Like	Forward (C)	5'- GGCACATCCAAGGCTGTAGCGCCAAAACGGGAGCAGGACTGGAGGAAGGCGTGGCATGGA TACTTAGCACTTGTGCCCCGactagtgtgagcaagg-3'
<i>Tb</i> ARL3 Like	Reverse (C)	5'- AGTGAGGCATGAGCGTGTCTATTGAAGTCTTCCCCAGGAAATTATTTGCTTTCAGCGCCTTA CAGACATCAGCTTCCAActattcctttgccctcggac-3'
<i>Tb</i> BARTL1	Forward (C)	5'- CAGTCACTGCTGGGAGATCTTGTTTGGTAAGTTCAGATTAGAAACGACATAATACTCCACATT TGTATTTGCAAAGGAActattcctttgccctcggac-3'
<i>Tb</i> BARTL1	Reverse (C)	5'- CAGTCACTGCTGGGAGATCTTGTTTGGTAAGTTCAGATTAGAAACGACATAATACTCCACATTT GTATTTGCAAAGGAActattcctttgccctcggac-3'
<i>Tb</i> HSP70	Forward (C)	5'- CCGGTGGTATGCCCGGAGGAATGGGTGGTGGGATGGGAGGCGCTGCGGCATCGTCCGG GCCTAAAGTCGAGGAGGTTGACactagtgtgagcaagg -3'
<i>Tb</i> HSP70	Reverse (C)	5'- ACGAAGAGAAAAATGTGGCGGCGGACCATCACCGTAGACAGTGAGACGACTGCAGCACCG GTCCGAAATACACCTGGGAACctattcctttgccctcggac -3'
<i>Tb</i> KMP11	Forward (C)	5'-TGCGCGAGCACTCAGAGCACTTCAAGGCCAAGTTTGCGGAACTCCTCGAGCAGCAGAAGA ATGCCAGTTCCCCGGAAAAactagtgtgagcaagg -3'
<i>Tb</i> KMP11	Reverse (C)	5'-ATCTTTTAAAAGATTTTTATAGTATTATCCCTAGGGACTCTAACATAGGAAAGAAAATGAG AGGTGTAGTAAATTTGTAActattcctttgccctcggac -3'

<i>Tb</i> MARP	Forward (N)	5'TATATCATTTTGAGCAAAACAGAAATCTTTTCTAGCTAAAGCATAAGCCAACCACATCGACG ACAAACAAGAACAAAGAatgcctttgtctcaagaag -3'
<i>Tb</i> MARP	Reverse (N)	5'-TTGTTTTTCTCTTCTCCGATATTTATGTTGTTTCATCTTGTGTTTTATATACTTGTGCTT TTTTTTTCGTACATcttgtacagctcgccatgc -3'
<i>Tb</i> CAP51	Forward (C)	5'- ATTCGGAGTTGATGCTGCTGAAGCATGACCTGCAGACCGTTCTTCACTATATCAGGGTCCG TGCGCGGGAGGCCGACAAActagtgtagcaagg -3'
<i>Tb</i> CAP51	Reverse (C)	5'- CCTTAACGTAATTAACCTCATACATAAAACGAAAGGATATATTTGGCGATGAAAACTGAA CATGGAGCGTAATACGATCctattcctttgccctcgac -3'
<i>Tb</i> Protein Kinase	Forward (C)	5'- GCGGAAACGACCTATTCCAAAGGCCACAGCAGCAACAGCCACAGCAACCAGAGAAGGAC CCCTTCGCCAGTCTCTTCAAGactagtgtagcaagg -3'
<i>Tb</i> Protein Kinase	Reverse (C)	5'- CCAACTAACCCCCCTTTTTTTTGGGGGGGGGGTGACGTATGATGCTGTTCTGTGCCAG TATTGCAGTCATGTAAACActattcctttgccctcgac -3'
<i>Tbp</i> 25 $\alpha$	Forward (C)	5'-GTAACGCCGGATCCCTTTCAGGGGTTGTCGACCGCCGTGTGGATCAGGTCGATGTGCGTGG CACAACCTCCTCGCAAAAGactagtgtagcaagg -3'
<i>Tbp</i> 25 $\alpha$	Reverse (C)	5'-CACCTACAAGTACCACAAATATATATATATATATATATATATATATCAAATGAATATAACATATTACAA TTTCTTTTGTtattcctttgccctcgac -3'
<i>Tb</i> TCP-1- $\varepsilon$	Forward (C)	5'-AACTCCGGCTCGCCACGCAGGTTGTGAAAATGATTTTGAAGGTTGATGACGTTATCGTAACAAAGCC CGAAGAGGAGGAAactagtgtagcaagg -3'

<i>Tb</i> TCP-1-ε	Reverse (C)	5'- TCGAAGTGTATATGTTCTGCGCATGACCCCCTCCCGCTCCAACATAGCGTTCCATCCCGTACCTCC TGCAACAGCACATctattcctttgccctcggac -3'
<i>Tb</i> VHS	Forward (C)	5'-TTTCTGAAGAAAGAAACCGGAAACCGGTTAAGCACGTTTATACCGATTTTGAAGACCTTTTGGCGCG ACAGGCACCAAAGactagtgtgagcaagg -3'
<i>Tb</i> VHS	Reverse (C)	5'- ACGCACGCACGCAAATCTACCCATGCTAATATCATGTATACTTGGTGTACTACCCACCCAGTGGA GAGTTATTGCTAACActattcctttgccctcggac -3'
<i>Tb</i> BBP590	Forward (C)	5'- CGATTCGCAGGGTGGTAATACCAACACAAGAGCAGTTGGGGACGCCGTCTCCGAAATACG ATTTCTGTACCCTAGGGAAactagtgtgagcaagg -3'
<i>Tb</i> BBP590	Reverse (C)	5'-AACGCCGAGCTCTTTTTTTTTTAAAAAACCAAACCTCAGCGAGCTAACAAGCTTGGAACAC ACATATATGTGTACACAActattcctttgccctcggac -3'
<i>Tb</i> BBP590	Forward (N)	5'- TTGTTGGCGTTACTTCTTTGTTATTACGGTTGCTGTTGTGGCAGACGCACATTTTTTGACATT ACAACTGCTGCACGTCGatgcctttgtctcaagaag -3'
<i>Tb</i> BBP590	Reverse (N)	5'-ATTTCACTCCTGATTGCGAGGAGATCCCCTTCATCAGAGTTGTACAGGTCCATCACAGGTCC CTGGGGAGCGAAGGGCATcttgtacagctcgtccatgc -3'

Capital letters: homology to appropriate region of gene locus to be targeted; lowercase letters: homology to pPOT vector

**Table 2.9 Sequences of oligonucleotide primers used to generate constructs for RNAi knockdown experiments.**

<b>Primer ID</b>	<b>Direction</b>	<b>Nucleotide sequence (For RNAi)</b>	<b>Restriction site</b>
<i>Tb</i> ARL3 Like	Forward	5'- GCG <u>gqatcc</u> TGAGCCGCGTGTACTGATAG -3'	BamHI
<i>Tb</i> ARL3 Like	Reverse	5'- GCG <u>ctcgaq</u> TTATTTGCTTTCAGCGCCTT -3'	XhoI
<i>Tb</i> BARTL1	Forward	5'- GCG <u>gqatcc</u> TTTACAGTTCACGCATTCGC -3'	BamHI
<i>Tb</i> BARTL1	Reverse	5'- GCG <u>ctcgaq</u> CCTGTAGCTGCAGATCCTCC -3'	XhoI
<i>Tb</i> HSP70	Forward	5'- TAT <u>gqatcc</u> ATCGAGATCGATGCACTCTTCG-3'	BamHI
<i>Tb</i> HSP70	Reverse	5'- GCG <u>ctcgaq</u> AATACCATTAGCGTCTAGGTCTG -3'	XhoI
<i>Tb</i> KMP11	Forward	5'- GCG <u>gqatcc</u> AGATCGTGCTTTCAAGAGTA -3'	BamHI
<i>Tb</i> KMP11	Reverse	5'- GCG <u>atcgac</u> GCATTGCGCTAGAAAAGTGG -3'	Sall
<i>Tb</i> MARP	Forward	5'- GCG <u>gqatcc</u> GGCCACTTTAGGACAACTGC -3'	BamHI
<i>Tb</i> MARP	Reverse	5'- GCG <u>ctcgag</u> CGCTGCTTACACCTTGCATA -3'	XhoI
<i>Tb</i> CAP51	Forward	5'- GCG <u>gqatcc</u> GGAAGAAACCGAGGAGCAGA -3'	BamHI
<i>Tb</i> CAP51	Reverse	5'- GCG <u>ctcgaq</u> CCTTCACGTGAGTACCACGGCG -3'	XhoI
<i>Tb</i> Protein Kinase	Forward	5'- TAT <u>gqatcc</u> TTCTGCTTCTCGCAGACTGAGC -3'	BamHI
<i>Tb</i> Protein Kinase	Reverse	5'- GCG <u>ctcgaq</u> AAAGAGGTCATCCGTTGTTGG -3'	XhoI

<i>Tbp25α</i>	Forward	5'- GCG <u>ggatcc</u> TGGAAGCTGTCTTCTACGCA -3'	BamHI
<i>Tbp25α</i>	Reverse	5'- GCG <u>ctcgaag</u> CACATCGACCTGATCCACAC -3'	XhoI
<i>TbTCP-1-ε</i>	Forward	5'- GCG <u>ggatcc</u> TCTCGCATCTCTGAGGGGT -3'	BamHI
<i>TbTCP-1-ε</i>	Reverse	5'- GCG <u>ctcgaag</u> TGCGGAAGTATTCCTGCTCT -3'	XhoI
<i>TbVHS</i>	Forward	5'- GCG <u>ggatcc</u> TATTGTGCTGGAGTCGCTTG -3'	BamHI
<i>TbVHS</i>	Reverse	5'- GCG <u>ctcgaag</u> TCCTCGTCCGTTGGTAAAAC -3'	XhoI
<i>TbBBP590</i>	Forward	5'- TAT <u>ggatcc</u> AAATTGTCGAGTCTGCGATCC -3'	BamHI
<i>TbBBP590</i>	Reverse	5'- GCG <u>ctcgaag</u> CGTCAGCATAAAGCGTTGAAGC -3'	XhoI

## 2.7 PCR amplification

For standard PCR procedure (*i.e.* for RNAi knockdown sequence amplification), 1  $\mu$ L of 10  $\mu$ M forward primer, 1  $\mu$ L of 10  $\mu$ M reverse primer, 3  $\mu$ L of gDNA template (50-100 ng) and 7.5  $\mu$ L of Milli-Q were added into Reddymix® PCR Master Mix (ThermoPrime) to make up final volume of 25  $\mu$ L (primers are shown in Table 2.9).

Number of Cycles	Stages	Temperature (°C)	Time
1	Initial denaturation	94	5 mins
25-29	Denaturation	94	30s
	Annealing	53	30s
	Elongation	72	1 min / 1 Kb
1	Extension	72	5 mins
	Hold	4	

**Table 2.10 Standard PCR program.**

For high fidelity PCR (*i.e.* for pPOT tagging), 2.5  $\mu$ L of 10  $\mu$ M forward and reverse primer, 3  $\mu$ L of pPOT vector template (25 ng/ $\mu$ L), 1  $\mu$ L of dNTPs (10 mM), 5  $\mu$ L of buffer with 15 mM MgCl<sub>2</sub> (10x), 1  $\mu$ L Expand™ Taq DNA Polymerase (2.5 U) and 33  $\mu$ L Milli-Q, which to make up final volume of 50  $\mu$ L (primers are shown in Table 2.8).

Number of Cycles	Stages	Temperature (°C)	Time
1	Initial denaturation	94	5 mins
29	Denaturation	94	30s
	Annealing	55	30s
	Elongation	72	1 min / 1 Kb
1	Extension	72	2 mins
	Hold	4	

**Table 2.11 High fidelity PCR program.**

## 2.8 Purification of PCR products

PCR products were purified using the GeneJET™ PCR purification kit protocol (see manufacturer's protocols), eluted in 50µl of filtered Elution Buffer and stored at -20°C. DNA concentration was quantified by NanoDrop 2000c/2000 UV-Vis Spectrophotometer (Thermo Scientific).

## 2.9 DNA Ligations

A ligation calculator (in silico) was used to determine the amount of vector or PCR insert required to achieve an insert:vector ratio of 3:1.

Ligation reaction:

50ng (1µl) of vector

Purified PCR products (X µL)

5µl x2 Rapid Ligase Buffer or 1µl of x 10 Rapid Ligase Buffer

1µl of T4 DNA ligase

Milli-Q (X µL) - to make up final total volume to 10 µL.

Ligations were left at room temperature for one hour then incubated at 16°C overnight before transforming into competent cells.

## 2.10 Transformation into Competent *E. coli* XL1-Blue cells

For each transformation, 1-3µl ligation reaction was added into 50µl aliquot of *E. coli* XL1-Blue subcloning-Grade competent cells. The cells were incubated on ice for a further 30 minutes followed by heat-shock at 42° for 45 seconds. Transformants were further incubated on ice for 2 minutes before the addition of 800µl of LB (pre-warmed to 37°C) and incubated at 37°C for one hour in an orbital shaker. Post incubation the culture was centrifuged and the bacterial pellet re-suspended in 200µl LB, and spread

onto an LB agar plate containing appropriate selective antibiotics. Plates were incubated at 37°C overnight.

## **2.11 Plasmid purification using a Miniprep kit**

Single *E. coli* (XL1-Blue) colonies were carefully picked from selective agar plates and inoculated into 3mL of LB media containing appropriate antibiotics and incubated overnight at 37°C in orbital shaker at 250-300 rpm. The bacterial cultures were centrifuged and plasmid DNA isolated using a Plasmid Miniprep Kit. Purified plasmid DNA was eluted in 50µl Elution Buffer then stored at -20°C.

## **2.12 Gel extraction and purification of DNA**

Plasmid DNA was digested with appropriate restriction enzymes overnight at 37°C. DNA was resolved on agarose gels as described in 2.13 and DNA visualised under the UVP Dual-Intensity Trans-illuminator to allow the excision of the desired band. Extracted DNA was purified using GeneJET Gel Extraction kit. Purified DNA was eluted in 50µl Elution Buffer and stored at -20°C.

## **2.13 Restriction enzyme digests**

Plasmid DNA was digested with appropriate restriction enzymes (Roche) and incubated at 37°C for 1 hour.

## **2.14 Agarose gel electrophoresis**

Gel electrophoresis was used to analyse PCR products and digested plasmid DNA. DNA samples were prepared by adding one volume of 6x loading buffer to five volumes of DNA. A 0.8% agarose gel was cast using 1xTAE and Gel Red™ (1:10,000 dilution). The agarose gel was run in submarine gel electrophoresis tank in 1xTAE loading buffer. The voltage and time was set at 80V and 45 mins. All samples were visualised and imaged using Image Lab™ software 4.0 Gel Red protocol (BioRad).



## **2.15 Preparation of plasmid DNA for transfection into *T. brucei***

Purified plasmid DNA was digested/linearised with NotI overnight at 37°C. The linear plasmid was ethanol precipitated by adding 1/10<sup>th</sup> volume of 3M sodium acetate (pH5.2) and 2.5X volume 100% ethanol. DNA was precipitated at -80°C overnight. Precipitated DNA was centrifuged at 22065 x *g* at 4°C for 45 minutes. The supernatant removed, the pellet washed in 250µl of 70% ethanol and centrifugation repeated. 70% ethanol was removed inside the Class II Microbiology safety cabinet to let residual ethanol evaporate. The dry DNA pellet was re-suspended in 50µl of sterile filtered Elution Buffer. For pPOT tagging, PCR reactions were purified and directly transfected into *T. brucei* (Dean et al., 2015).

## **2.16 *T. brucei* cell culture**

Procyclic form *T. brucei* cells were cultured in SMD-79 medium, supplemented with 10% foetal calf serum and Hemin (0.008mM) at 28°C (Adam, 2001).

## **2.17 Transfection of DNA into *T. brucei***

3x10<sup>7</sup> cells ml<sup>-1</sup> were required for each transfection. An appropriate volume of cell culture was centrifuged, re-suspended in ZMG buffer (0.5mL) and transferred to electroporation cuvettes (Bio-Rad 0.4 mm size) prior to addition of 5-10 ng of linearised DNA. Cells were electroporated using a BTX Electro Square Porator ECM830 (3 x 100 µs pulses at 1700 V with 200 ms intervals). Cells were transferred from the cuvette into 10ml of SMD-79 media (supplemented with 10% foetal calf serum and 0.008mM Hemin) and left to recover overnight. Following overnight recovery cells were diluted to 5x10<sup>5</sup> cells ml<sup>-1</sup> in SMD-79 media and selected using appropriate drugs.

## 2.18 Glycerol stock for *T. brucei* cultures

For all transformed cell lines (once they had become healthy and a reasonable density), 4.5mL of cell culture and 0.5 mL of 100% autoclaved glycerol were mixed in a falcon tube, and 1 ml aliquoted into a cryovial. 1 mL aliquots were frozen first at -80°C overnight, and then transferred to liquid nitrogen for long term storage.

## 2.19 Sample preparation for mass spectrometry (BioID-SILAC)

### 2.19.1 Culturing of *T. brucei* in SILAC media

SDM-79-SILAC liquid media was a custom synthesis from Cassion Labs, USA (See Table 2.1). SDM-79-SILAC media lacks the essential amino acids L-arginine and L-lysine. Isotopically labelled L-arginine and L-lysine were made up according to Table 2.2 in Milli-Q and appropriate labels (light-, medium- or heavy-) added to the media.

Different *T. brucei* cell lines were cultured in normal SDM-79 media first, and diluted to a starting density of  $1 \times 10^6$  cells  $\text{ml}^{-1}$  in SDM-79-SILAC media in a final volume of 10 ml. The appropriate SILAC labels were added into designated flasks depending on the experiments. Parental cell line (*T. brucei* PTPs) were always cultured using light-labels (K0R0), cells expressing *TbRP2* <sup>$\Delta 134-463$</sup> ::myc::BirA\* and cells expressing *TbRP2*::myc::BirA\* were cultured in medium-labels (K4R6) and heavy-labels (K8R10) respectively. Flasks were kept at 28°C and cell counts carried out regularly (using a haemocytometer; Neubauer) to ensure cells had undergone at least seven cell divisions (the minimum number of divisions required to ensure all proteins synthesised within a cell had incorporated the isotopically-labelled amino acids) before being scaled up to a final volume of 40 ml in SDM-79-SILAC (plus appropriate labels) with a starting density of  $2 \times 10^6$  cells  $\text{ml}^{-1}$ . After cells were transferred into separate flasks with SDM-79-SILAC and appropriate labels, doxycycline was added to induce expression of the BirA\* fusion protein. After 24-hour incubation at 28°C, excess biotin (final concentration 50  $\mu\text{M}$ ) was added to all cultures, and cells incubated for a

further 24 hours. For the experiments,  $1 \times 10^9$  cells were needed ( $\sim 3.33 \times 10^8$  cells per cell line). It is critical to have equal number of cells per cell line, otherwise this affect the SILAC-MS results. The appropriate volume of each cell line (to give  $3.33 \times 10^8$  cells) was centrifuged at  $1000 \times g$  for 10 min at  $4^\circ\text{C}$  and the supernatant removed. Cell pellets were washed three times in filtered 1xPBS (40mL), re-suspended in 1ml and transferred into an Eppendorf tube. Cells were lysed in 1ml of ice cold filtered Milli-Q in the presence of an inhibitor cocktail (20  $\mu\text{l}$  of the Halt Protease Inhibitor single-use cocktail, 0.1 mM TLCK, 1 mM benzamidine and 1 mM PMSF) and incubated at room temperature for 10 min. Lysed cells were divided into aliquots of  $5 \times 10^8$  cell equivalents and stored at  $-80^\circ\text{C}$ .

- **Cytoskeleton sample**

Cells were transferred into fresh Eppendorf tubes, centrifuged  $1800 \times g$  for 3 mins in bench-top microfuge and excess PBS removed. Cells were extracted in 1ml of extraction buffer per  $1 \times 10^9$  cells, with protease inhibitors (see above). Samples were left on an orbital shaker for 15 mins at room temperature. Samples were centrifuged at  $8502 \times g$  in bench-top microfuge for 10 min and the supernatant transferred into a fresh Eppendorf tube. BOTH the pellet and supernatant were stored at  $-80^\circ\text{C}$  for future analysis.

## **2.19.2 BioID Sample Preparation (on beads digest)**

*T. brucei* PTP cells and cells expressing *TbRP2* <sup>$\Delta 134-463$</sup> ::myc::BirA\* and *TbRP2*::myc::BirA\* on a PTP background were cultured in SDM-79-SILAC (light-, medium- or heavy- labels), and cells lysed and extracted using the protocol described in Section 2.19.1. The cells (pellet or solution) were defrosted on ice, diluted with solubilisation buffer at a 1:1 ratio (or 1ml solubilisation buffer for the pellet sample) and left on the orbital shaker for 30 mins at room temperature (and vortexed every 10 mins until they became clear). The samples were centrifuged at  $15,521 \times g$  (Beckman Coulter, Allegra™ X-22R centrifuge) at  $4^\circ\text{C}$  for 10 mins, and the supernatant transferred to a fresh LoBind Eppendorf tube (Sigma-Aldrich) and kept on ice. Streptavidin-coated magnetic beads (100  $\mu\text{l}$  per column) were washed with filtered 1x

PBS (1mL x 3 time) in LoBind Eppendorf tube on a magnetic rack. Washed Streptavidin-coated magnetic beads were mixed with the sample supernatant and incubated on a 360-degree rotating wheel at 4°C for 4 hours (to ensure that biotinylated molecules adhered to streptavidin immobilised on the magnetic beads). After 4 hours incubation, the LoBind Eppendorf tubes were placed on a magnetic rack (for 2-3 mins) and the supernatant removed. The streptavidin-coated magnetic beads were then re-suspended in 1 ml of IAA solution, which contained iodoacetamide for the reductive alkylation of free cysteine residues. The LoBind Eppendorf tube was wrapped in tin foil to exclude light and incubated on a 360-degree rotating wheel at room temperature for 20 min. The LoBind Eppendorf tube was again placed on the magnetic rack and the supernatant removed. The beads were washed with filtered 1ml of 1xPBS (x 5).

#### - **Digestion**

The reduced and alkylated beads/sample was re-suspended in 50µL of 50mM ABC, followed by 1µl trypsin. The sample was vortexed, the LoBind Eppendorf tube sealed with parafilm and then incubated at 37°C for at least 18 hours.

#### - **Peptide desalting**

After the beads/sample had been incubated overnight (the tube was given a quick pulse to break up any clumps). The magnetic column rack was used to separate the magnetic beads from the supernatant (containing digested peptides). The supernatant was transferred into a fresh LoBind Eppendorf tube. The beads were then washed with 100µL of 50mM ABC solution and the resulting supernatant separated from the magnetic beads using the magnetic rack and the supernatant combined with the previous supernatant in the LoBind Eppendorf tube. Finally, the beads were washed with 100µL of 0.5 M NaCl and the supernatant separated from beads as previously described before being combined with the two previous supernatants (250µL of digested peptides in total).

A C18 microspin column (Harvard Apparatus) was washed with 100 µL of methanol and centrifuged at 200g for 1 min and the flow-through discarded. The microspin column was then washed 3 x 100 µL with solution containing 0.1% TFA and 70% MeCN

and centrifuged at 200g for 1 min and the flow-through discarded. The column was then washed 3 x 100  $\mu$ L with solution contains 0.1% TFA and centrifuged at 200 x *g* for 1 min and the flow-through discarded. TFA was then added to the LoBind Eppendorf tube containing the digested peptides (0.1% TFA final concentration). The peptides were then transferred to the microspin column for desalting. The microspin column was centrifuged at 200 x *g* for 1 min and the flow-through discarded. The microspin column was washed 3 x 100 $\mu$ L with 0.1% TFA solution and centrifuged at 200 x *g* for 1 min. The sample was eluted with 2x 100  $\mu$ L of 0.1% TFA, 70% MeCN solution, and centrifuged at 200 x *g* for 1 min and collected into a new LoBind Eppendorf tube. Finally, 200  $\mu$ L of sample was diluted with 500  $\mu$ L of filtered HPLC water and stored at -80°C until the sample was frozen. The frozen samples were then lyophilised (Freeze-dryer, ALPHA 2-4 LD plus, Martin Christ) to await further analysis by mass spectrometry.

### **2.19.3 BioID Sample Preparation (in gel digest)**

Cells were harvested as described in 2.20.2. After the beads/sample was washed in 1X PBS, the beads/sample was heated in 50 $\mu$ L of Laemmli sample buffer at 95°C for 5 min. The beads/sample was centrifuged at 200g and the supernatant transferred to a fresh LoBind Eppendorf tube. The protein sample was run on a pre-cast 12% SDS-PAGE gel for 10 min at 200 V (the sample was allowed to migrate into the resolving gel for no more than 1-2 cm). The SDS-PAGE gel was stained with Instant Blue and de-stained using Milli-Q.

After the de-stain solution was removed, the gel was transferred onto a clean plastic tray and a small amount of Milli-Q added to prevent the gel from drying out. The region containing the stained protein was excised using a scalpel and the gel cut into 1x1 mm squares (taking care to ensure that the gel was not cut into pieces that were too small as these would clog pipette tips). Gel pieces were transferred into a fresh LoBind tube and centrifuged and removed all the liquid. 100 $\mu$ L of mixture of 100 mM ammonium bicarbonate and 100% acetonitrile (1:1, vol/vol) was added into the LoBind Eppendorf tube contains the gel pieces and incubated for 30 min with

occasional vortexing until the gel pieces were completely clear (removed all the liquid). 500µl of 100% acetonitrile was then added into the tube and incubated at room temperature (about 10 mins) with occasional vortexing, until gel pieces become white and shrank. The supernatant was discarded and about 30–50µl of the 10mM DTT in 100mM ammonium bicarbonate solution was added into the tube (enough to ensure gel pieces completely covered), and incubated at 56°C using air thermostat for 30 min. The tube was then cooled to room temperature (~ 22°C), followed by adding 500µl of 100% acetonitrile, the sample was incubated for a further 10 min and then all liquid removed. About 30–50µl of iodoacetamide solution (55mM iodoacetamide in 100mM ammonium bicarbonate) was added to the tube to cover the gel pieces, and incubated for 20 min at room temperature in the dark, before all liquid was removed. Finally, sufficient volume of 100% acetonitrile was added to shrink gel pieces for ~10 min and the all liquid removed.

Following the Sigma trypsin singles protocol (Sigma Aldrich):

5 µL Trypsin Solubilization Reagent was added to a Trypsin Singles vial, and the vial vortexed briefly to ensure trypsin is dissolved. 45 µL of the Trypsin Reaction Buffer was also added to the same vial and mixed well (final trypsin concentration is 20 mg/ml). 50µL of trypsin buffer was added to cover the dry gel pieces and incubated on ice or in the fridge for 30 mins. (NOTE - after 30 minutes the tube was examined to determine whether all the solution had been absorbed, and if it had not then more trypsin buffer was added as gel pieces should be completely covered with trypsin buffer). Gel pieces were incubated on ice for a further 90 minutes to saturate them with trypsin, and then 10–20µl of 10mM ammonium bicarbonate containing 10% (vol/vol) acetonitrile buffer was added to the tube to cover the gel pieces (gel pieces must be kept wet during enzymatic cleavage). The LoBind Eppendorf tube containing the gel pieces was sealed with Parafilm and incubated at 37°C overnight (NOTE - ensure that the volume added is recorded as this information is required).

After tryptic digestion, digested peptides were extracted from gel pieces using extraction buffer (1:2 (vol/vol) of 5% formic acid/acetonitrile). An approximate ratio of 1:2 between volumes of digest and extraction was achieved and the sample

incubated at 37°C in an orbital shaker for 15 minutes. Supernatant was collected into a new LoBind Eppendorf tube (extracted gel pieces were not discarded). The desalting procedure using a C18 microspin column as described in Section 2.19.2 was followed and eluted peptides diluted with 500µL of filtered HPLC water and stored overnight at -80°C. The frozen samples were then lyophilised (Freeze-dryer, ALPHA 2-4 LD plus, Martin Christ) for further analysis by mass spectrometry.

## **2.20 Liquid chromatography mass spectrometry**

All lyophilised samples were sent to the FingerPrints Proteomics Facility at the College of Life Sciences, University of Dundee, for analysis by liquid chromatography tandem mass spectrometry (LC-MS/MS).

## **2.21 Analysis of mass spectrometry data**

The methods and parameters for LC-MS/MS analysis was as described by Urbaniak et al., (Urbaniak et al., 2012). The raw data was processed using MaxQuant (Cox and Mann, 2008) version 1.5.2.8 which uses the Andromeda search engine (Cox et al., 2011). Proteins identified by LC-MS/MS were searched against the *T. brucei brucei* 927 annotated protein sequence database (Release 26, downloaded from TriTrypDB (Aslett et al., 2010); <http://www.tritryp.org>) and supplemented with frequently observed contaminants such as trypsin, BSA and human keratins. Search parameters specified a MS tolerance of 5 ppm, a MS/MS tolerance at 0.5 Da and full trypsin specificity, allowing for up to three missed cleavages. Carbamidomethylation of cysteine was set as a fixed modification and oxidation of methionines, N-terminal protein acetylation and N-pyroglutamate were allowed as variable modifications. Peptides were required to be at least 6 amino acids in length, and false discovery rates of 0.01 were calculated at the levels of peptides, proteins and modification sites based on the number of hits against the reversed sequence database. SILAC ratios were calculated using only peptides that could be uniquely mapped to a given protein. The output from MaxQuant was analysed using the Perseus framework.

## **2.22 Preparation of slides for fluorescence microscopy**

Approximately 1ml of cell culture was centrifuged at 1811 x *g* for 3 minutes in a bench top micro-centrifuge. The supernatant was discarded and the pellet washed with 1ml of 1xPBS. Centrifugation was repeated by adding ~800μL of 1xPBS at 1811 x *g* for 3 minutes. The cells were re-suspended and pipetted onto slides. After cells had settled, excess solution was removed. Whole cells were fixed in 3.7% paraformaldehyde (PFA) for 15minutes. To prepare detergent-extracted cytoskeletons, 1% PEME NP-40 (pH6.9) buffer was used, with cells extracted for 25 seconds before fixation in PFA. All slides were further fixed in 100% methanol at -20°C for a minimum for 10 minutes prior to storage in 100% methanol at -20°C.

## **2.23 Probing slides for fluorescence microscopy**

Slides were rehydrated in 1xPBS for 10 minutes, and then incubated in blocking buffer for one hour in a hydration chamber. Excess blocking buffer was removed and slides incubated in ~50μL of appropriate primary antibody for one hour. Slides were washed three times in wash buffer. The washed slides were incubated with the appropriate secondary antibody in the hydration chamber for further 40 minutes. Slides were washed three times in wash buffer again. Once excess wash buffer was removed, a single drop of Vectashield Mounting Medium with DAPI (Vectorlabs) was added on the slides before mounting. Sealed the edges of the slides by using nail varnish and stored the slides at 4°C in the fridge.



## 2.24 Fluorescence microscopy

Slides were imaged using an Applied Precision DeltaVision Deconvolution microscope and images processed using SoftWoRx software. All images were subsequently processed using Adobe Photoshop.

## 2.25 Preparation of *T. brucei* protein samples

$5 \times 10^6$  cells were required per lane for SDS-PAGE experiments. Trypanosome cells were centrifuged at  $9033 \times g$  for 10 minutes, and the supernatant discarded. The pellet was re-suspended in 1ml of 1xPBS and transferred into an Eppendorf tube. Re-suspended cells were centrifuged again for 3 minutes at  $9033 \times g$  and the supernatant removed. The cell pellet was re-suspended in 1xSDS loading buffer (pre-heated to  $100^\circ\text{C}$ ) and samples heated on a heat block at  $100^\circ\text{C}$  for further 5 minutes. Protein samples were stored at  $-80^\circ\text{C}$ .

## 2.26 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

10% SDS-PAGE gels were prepared according to standard protocols (Bruggerolle and Mignot, 2003). All protein samples in SDS loading buffer were heated at  $100^\circ\text{C}$  for 5 minutes.  $6\mu\text{l}$  of molecular weight marker ladder per lane (Precision Plus Protein™) and  $10\mu\text{l}$  of protein sample per lane ( $5 \times 10^6$  cell equivalents) were loaded in each well. Proteins were separated at 200 volts using a Mini-PROTEAN®3 Cell at 200 volts for one hour.

## 2.27 Immunoblotting

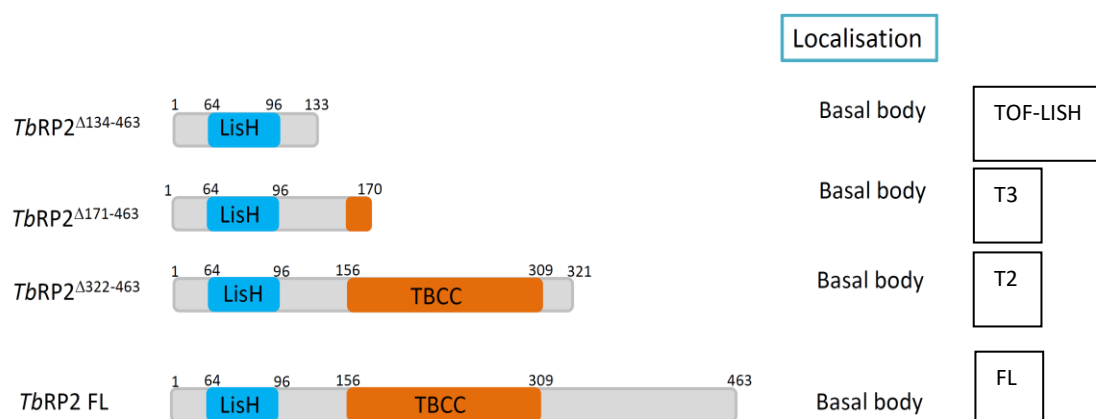
Amersham Hybond-P PVDF Membranes (GEHealthcare) were incubated in 100% methanol for 10 seconds, and then incubated first in Milli-Q for 5 minutes, and then western blot transfer buffer for a minimum of 10 minutes. Blotting paper and sponge pads were also pre-incubated in western blot transfer buffer. The blotting papers, pads, membrane and SDS-PAGE gel were assembled in correct order and placed into

a BioRad blotting cassette. The SDS-PAGE gel was transferred at 100 volts for one hour. After transfer was complete the Amersham Hybond-P PVDF membrane was stored at 4°C. Membranes were re-permeabilised in 100% methanol for 10 seconds, in Milli-Q for 5 minutes, and then washed twice in 1xPBS for 5 minutes. Blots were blocked by incubation in blocking buffer for 1 hour. Membranes were then incubated in appropriate primary antibody for one hour. After one hour incubation, primary antibody solution was removed and the membrane washed 3 x 10 minutes in western blot washing buffer. Membranes were then incubated with the appropriate secondary antibody for one hour at the recommended dilution. Membranes were then washed 3x 10 minutes in washing buffer before being incubated with Immun-Star™ WesternC™ Chemiluminescent Detection reagent for 5 minutes. The membrane was then imaged using a Chemi DOC XRS imager (BioRad) and images collected using a CCD imaging system.

# Chapter 3 Identification of putative TbRP2 interacting/near neighbour proteins using proximity-dependent biotinylation

## 3.1 Introduction

Previous studies have demonstrated that *TbRP2* localises to transitional fibres radiating from the mature basal body. RNAi mediated ablation of *TbRP2* results in the assembly of short and structurally defective flagella indicating that *TbRP2* performs critical flagellum assembly-related functions (Stephan et al., 2007). However, the precise role of *TbRP2* in flagellum biogenesis is unclear (Andre et al., 2014; Stephan et al., 2007). To address this question, I set out to identify proteins that interacted with, or were close to, *TbRP2* in trypanosome cells. By identifying *TbRP2*-proximal proteins, it was anticipated that further insight could be provided into: (i) trypanosome GTPases that were the molecular clients of *TbRP2* GAP activity, (ii) cytoskeletal proteins that contribute to the overall architecture of the mature basal body, and/or transitional fibers, and (iii) proteins responsible for transporting *TbRP2* to the mature basal body. Previous work from our laboratory (Andre et al., 2014), has shown that deletion constructs that (1) only contained the N-terminal TOF-LisH motif (and lacked the TBBC-like domain); (2) contained the N-terminal TOF-LisH motif and part of the TBCC domain; (3) contained the N-terminal TOF-LisH motif, full length of TBCC domain and without C-terminus (Figure 3.1) were all targeted to and tethered at the mature basal body. Therefore, the approach taken in this study was to express full length, truncated and mutant versions of the *TbRP2* protein fused to a promiscuous bacterial biotin ligase; an approach referred to as proximity dependent biotinylation (BioID). The intention being to investigate if interaction proteins/near neighbours profiles identified by BioID approaches differed between full length (FL) and truncated versions of *TbRP2*.

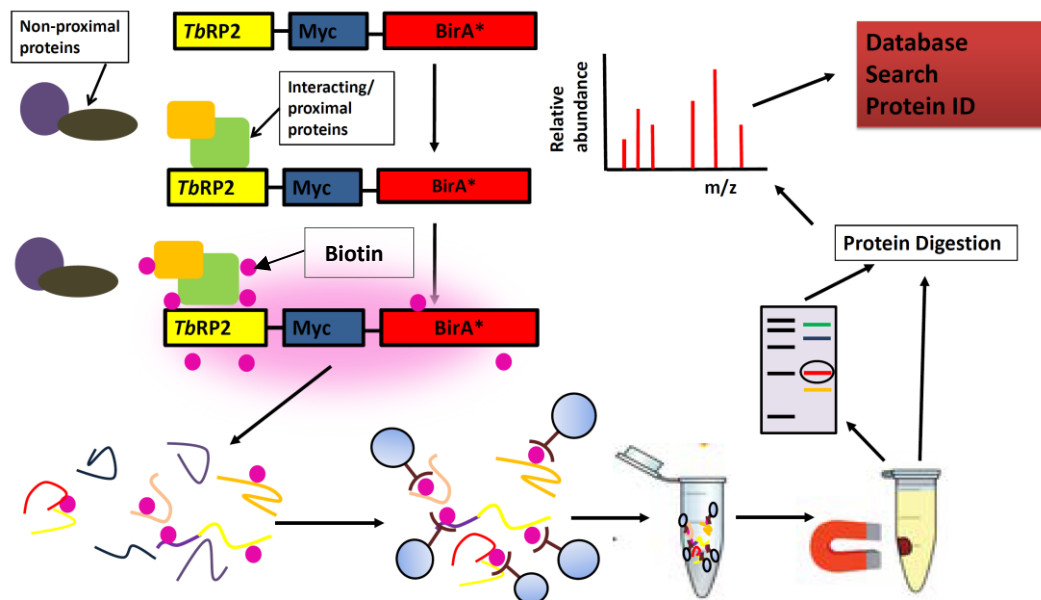


**Fig 3.1 showing schematic architectures and localization of C-terminally deleted proteins generated by** Andre et al., 2014. These constructs were further cloned into the pDEX\_myc-BirA\* vector.

## 3.2 BioID in theory

Identifying potential interacting proteins/near neighbours can be very informative in investigating proteins of unknown function, i.e. to interrogate function in the context of an interaction network. Additionally, mapping protein-protein interactions can provide key roles in understanding human disease processes, as major biological processes are mediated through protein interactions (Novarino et al., 2011). Proximity-dependent biotin identification (BioID) technique is a relatively newly developed method, which allows a forward screen to identify novel binding or nearby neighbour partner proteins (Roux et al., 2012). In *T. brucei*, BioID has been successfully employed to identify bilobe components interacting with the bilobe protein MORN1 (Morriswood et al., 2013; Morriswood et al., 2009); proteins that interact with Polo-like kinase (*TbPLK*) (which localises to different places during the cell cycle; (McAllaster et al., 2015), and as well as new transition zone proteins (Dean et al., 2015).

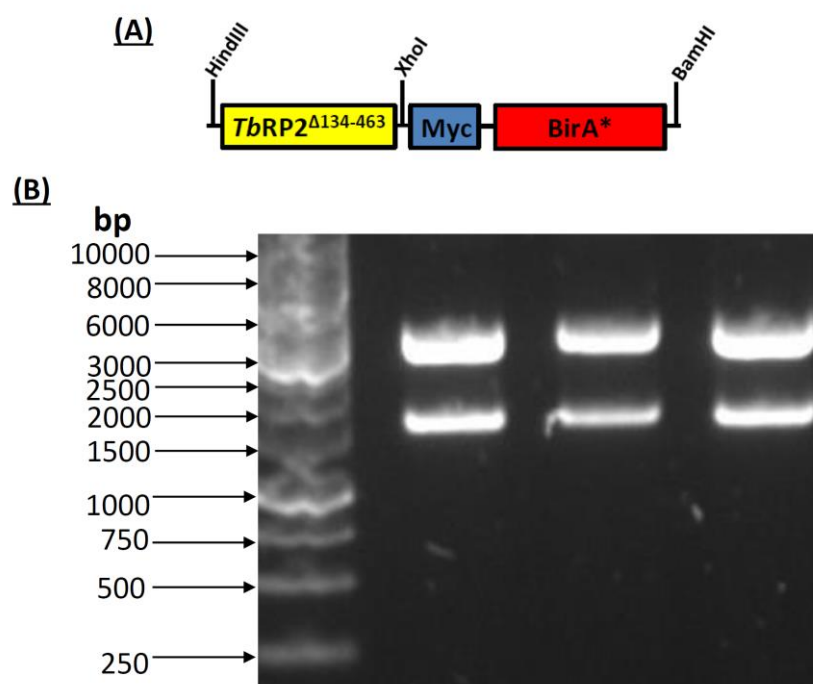
The principle of BioID involves tagging a protein of interest with a modified form of the bacterial biotin ligase called BirA\*, which can generate a highly reactive AMP-biotin intermediate. When this active form of AMP-biotin is released into cytoplasm it will react rapidly with lysine residues on a protein substrate, leading to biotinylation of proximity proteins (within a 10-12nm radius) (Roux et al., 2012). Potential interacting and/or proximity proteins within that biotinylation radius will get biotinylated; those are not close enough proteins will not. The biotinylated proteins can be purified using streptavidin beads and analysed by Mass Spectrometry (Fig 3.2).



**Figure 3.2 Schematic showing the application of the BioID method.** Here I used a PTP cell line to inducibly express *TbRP2::myc::BirA\**. Cells were induced to express *TbRP2::myc::BirA\** with doxycycline for 24 hours and then 50  $\mu$ M biotin added for a further 24h induction period. Cells were harvested, lysed and denatured under stringent conditions. Biotinylated candidate proteins were affinity purified using streptavidin-conjugated beads, and samples analysed by mass spectrometry or immunoblot analysis.

### 3.3 Generation of plasmids to allow expression of *TbRP2::myc::BirA\** fusion proteins

My aims here were to generate constructs which allow *T. brucei* cells to express *TbRP2*<sup>Δ134-463</sup> (TOF-LISH), *TbRP2*<sup>Δ171-463</sup> (T3), *TbRP2*<sup>Δ322-463</sup> (T2) as well as *TbRP2* full length (FL) BioID in an inducible fashion. The synthetic products of *TbRP2*<sup>Δ134-463::Myc::BirA\*</sup> were sub-cloned into pEK-K4 Vector, after transformation and DNA extraction steps. Plasmid DNA was prepared and subjected to restriction mapping analysis. Figure 3.3 shows plasmid DNA digested with HindIII and BamHI to identify plasmids that contained appropriate DNA inserts. Expected DNA fragments (~1500 bp) and vector (~2500bp) can be seen on the gel.

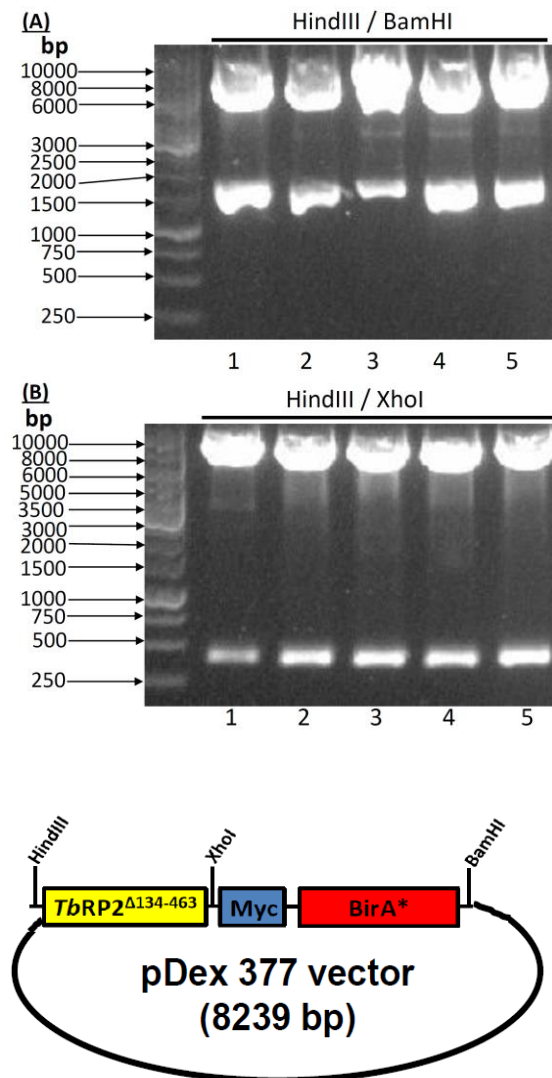


**Figure 3.3 (A) Schematic of DNA construct of *TbRP2*<sup>Δ134-463</sup>::myc::BirA\* in pEK-K4 vector. (B) Agarose gel showing DNA constructs digested with HindIII/BamHI; expected product size ~1500bp and vector 3000bp.**

The *TbRP2*<sup>Δ134-463</sup>::myc::BirA\* insert was sub-cloned into the *T. brucei* pDex377 expression vector. The pDex377 vector contains the tetracycline-inducible EP1 promoter to control transgene expression, and allows the transgenes to be stably integrated into mini-chromosomes and intermediate-sized chromosomes (177

bp repeat) (Kelly et al., 2007). Restriction mapping results and schematic of the DNA construct of *TbRP2*<sup>Δ134-463</sup>::myc::BirA\* in pDex377 are shown in Fig 3.4. DNA constructs were digested with HindIII/BamHI (A), which released the *TbRP2*<sup>Δ134-463</sup>::myc::BirA\* fragment (~1500bp) and pDex-377 vector (~8000bp). When digested with HindIII/XhoI (B) the *TbRP2*<sup>Δ134-463</sup> fragment (~400bp) was released. Positive results were seen in all five colonies.

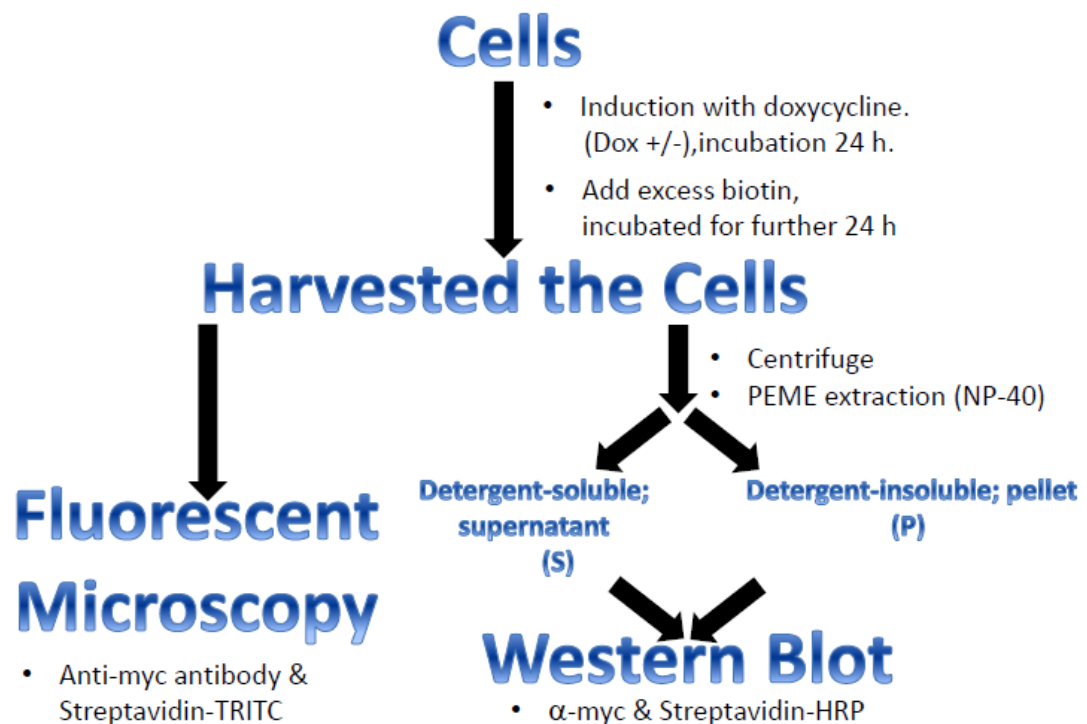
pDEX377-derived plasmids were linearised with NotI prior to transfection into procyclic *T. brucei* (SmOxP9). The DNA fragment encoding *TbRP2*<sup>Δ134-463</sup> was replaced by *TbRP2*<sup>Δ171-463</sup>, *TbRP2*<sup>Δ322-463</sup> and Full Length *TbRP2* for BioID experiments.



**Fig 3.4. Gel analysis of restriction digested *TbRP2*<sup>Δ134-64</sup>::myc::BirA\* in pDex377 clones. (A)** DNA constructs were digested with HindIII/BamHI, product size ~ 1500bp. **(B)** DNA constructs were digested with HindIII/XhoI, product size ~ 400bp.

### 3.4 Localisation of *TbRP2::myc::BirA\** expressing protein in *T. brucei*

Before moving my experiments further, the intracellular localisation of the various *TbRP2* proteins were investigated by immunofluorescence microscopy to ensure that adding the BirA\* sequence did not affect the localisation of the *TbRP2* protein. A preliminary fractionation protocol was designed to accomplish a two stages purification (Fig. 3.5). Both detergent-soluble and detergent-insoluble samples were investigated, and cells expressing *TbRP2::myc::BirA\** and *TbRP2*<sup>Δ134-463</sup>::*myc::BirA\** fractions were compared to achieve the optimised experimental condition for further BiolD protocols.



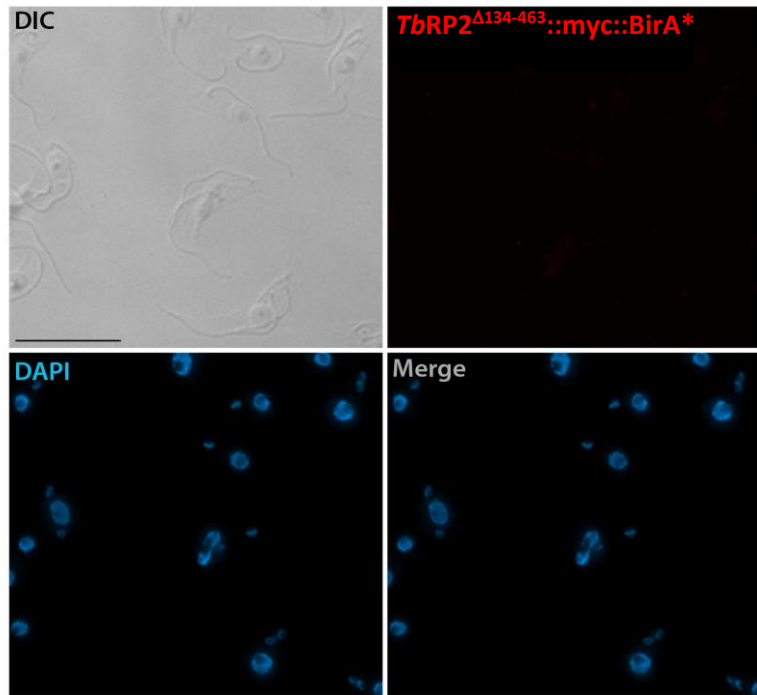
**Fig.3.5** The diagram shows the purification procedures to determine the localisation and expressions of *TbRP2::myc::BirA\**. Cells were grown in media with/without doxycycline for 24h, followed by incubation in excess biotin for a further 24h. Cells were harvested by centrifugation and incubated with 0.5% NP-40. Cells were further separated by centrifugation into detergent-soluble (S) and detergent-insoluble (P) fractions. Localisation and expression was analysed by immunofluorescence analysis (IFA) and immunoblotting.



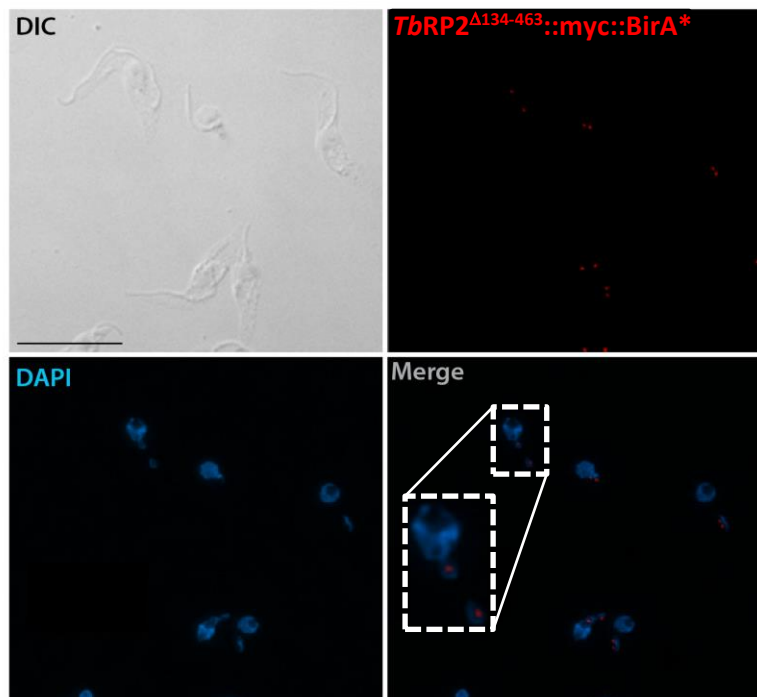
Indirect immunofluorescence analysis of detergent-extracted cytoskeletal cells from all four *TbRP2::myc::BirA\** cell lines are shown in Fig 3.6 to Fig 3.13. The expression and localisation of all four *TbRP2::myc::BirA\** proteins were analysed by probing cells with either anti-myc antibody or TRITC-conjugated streptavidin-- slides were imaged using Delta Vision (DV) microscopy. Fig 3.6 and Fig 3.7 show the *TbRP2<sup>Δ134-463</sup>::myc::BirA\** cells labelled with anti-myc and TRITC-conjugated streptavidin respectively. In the presence of doxycycline, a discrete signal of *TbRP2<sup>Δ134-463</sup>::myc::BirA\** is observed near/distal to the kinetoplasts (i.e. the expected basal body localisation). In the absence of doxycycline, no protein expression was detected with the anti-myc antibody. However, there was a trace of nuclear staining in cells labelled by streptavidin-TRITC, which might reflect naturally biotinylated proteins in the cells.

In Fig 3.8 and Fig 3.9, detergent-extracted cells expressing *TbRP2 FL::myc::BirA\** fusion was labelled with anti-myc antibody and streptavidin-TRITC. In the presence of *doxycycline*, the *TbRP2 FL::myc::BirA\** fusion protein can be clearly localised at the basal body by both the anti-myc antibody and TRITC-streptavidin. The localisation is again consistent with the published localisation of endogenous *TbRP2* (Stephan et al., 2007). In the absence of doxycycline, detergent-extracted cells showed no significant signal. Expression of either full length or the C-terminally truncated *TbRP2<sup>Δ134-463</sup>::myc::BirA\** proteins had no effect on cell morphology. However, *T. brucei* cells expressing *TbRP2<sup>Δ322-463</sup>::myc::BirA\** and *TbRP2<sup>Δ171-463</sup>::myc::BirA\**, appeared to have abnormal morphology after 24 hour of induction (Fig 3.10 - 3.13). The majority of cells appeared to have the kinetoplast positioned very close to the nucleus. A short new flagellum could also be observed in some of cells, which will lead to cell division defects, and cells with a 1 kinetoplast but no nucleus (i.e. a zoid) and 1 kinetoplast and 2 nuclei were also observed. Thus, even though these two truncated proteins correctly localised to the basal body following induction I decided not to use these two cell lines for BioID experiments due to the dominant negative phenotype that was induced.

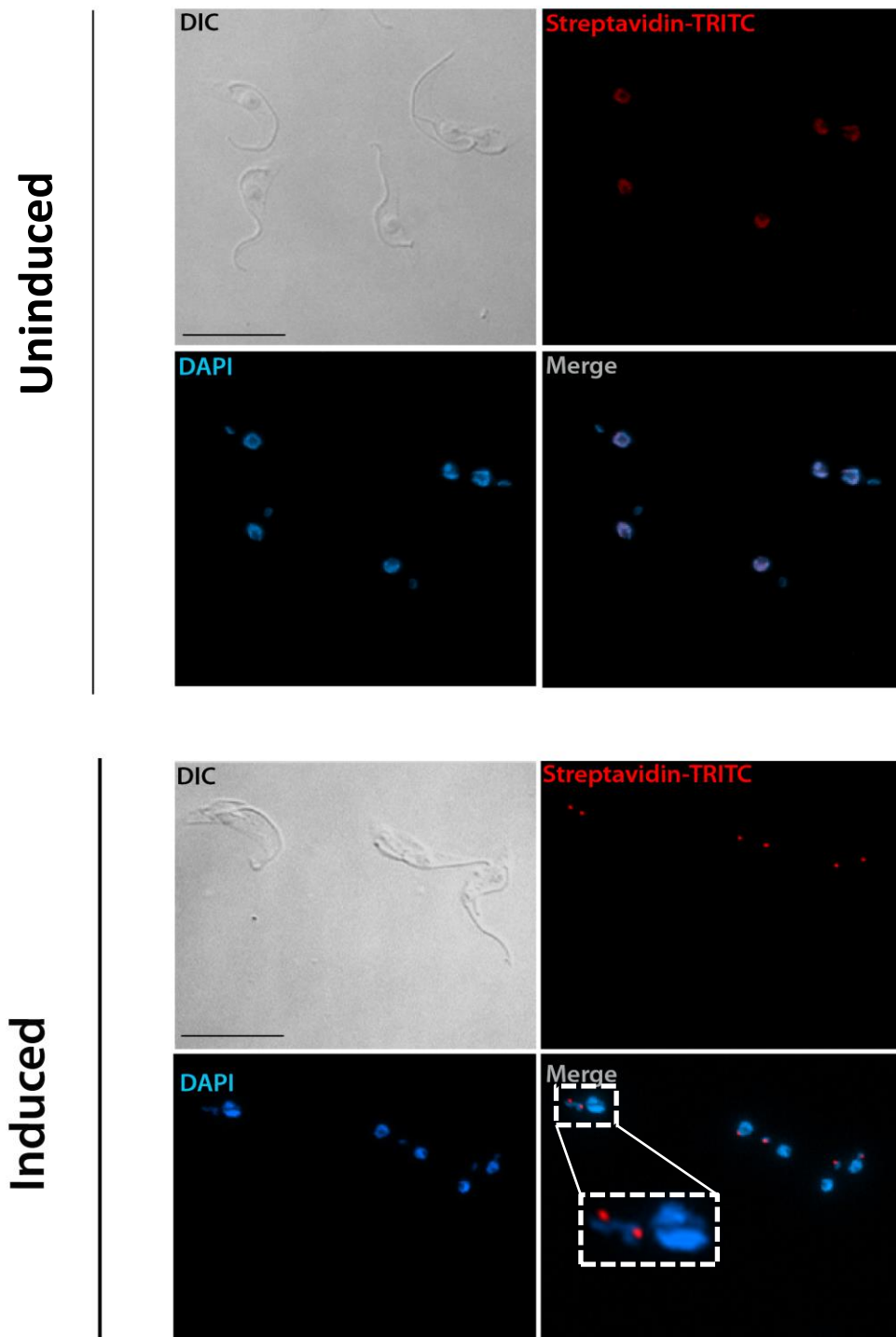
Uninduced



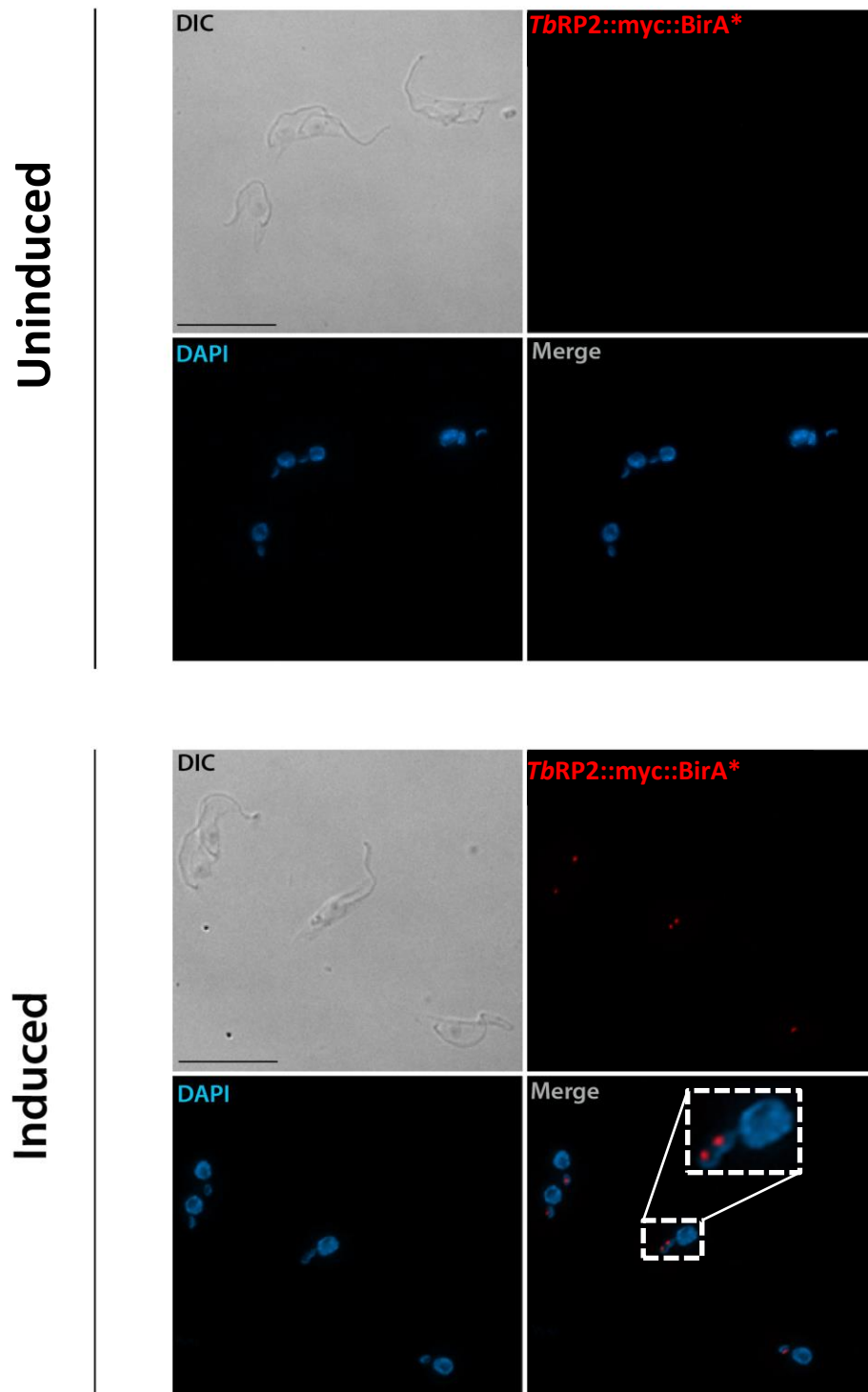
Induced



**Figure 3.6 Immunofluorescence analysis of a detergent extracted cells expressing *TbRP2*<sup>Δ134-463::myc::BirA\*</sup>.** Anti-myc antibodies were used to detect the *TbRP2*<sup>Δ134-463::myc::BirA\*</sup> protein. No expression can be detected in non-induced cells. In induced cells, fusion protein was clearly observed in the region of the basal body. Scale bars = 10 μm. Inset box shows the indicated region at higher magnification.

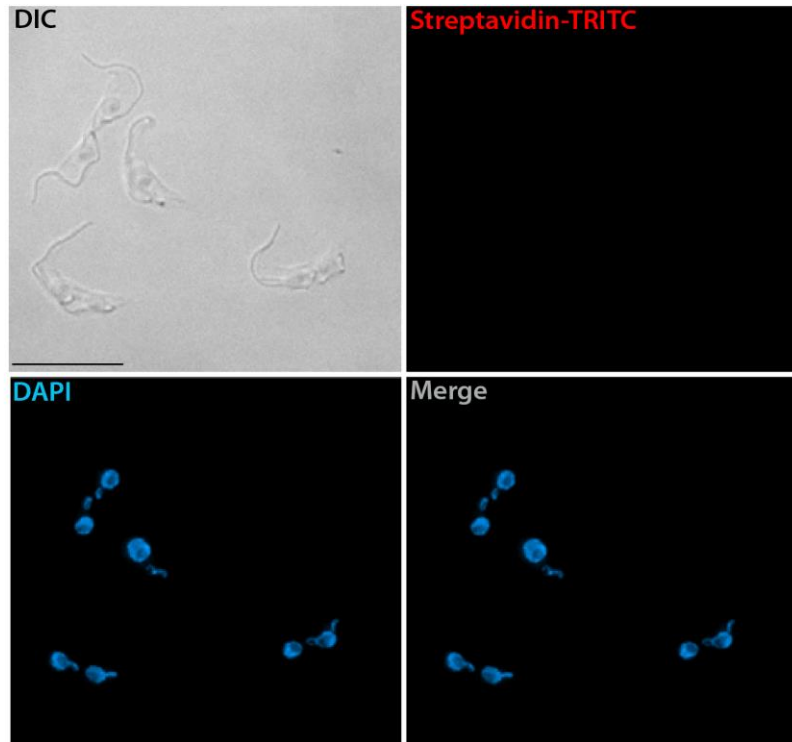


**Figure 3.7 Immunofluorescence analysis of detergent extracted cells expressing *TbRP2*<sup>Δ134-463::myc::BirA\*</sup> probed with TRITC-conjugated streptavidin.** Weak nuclear staining can be observed in non-induced cells. However, after induction, biotinylated proteins can only be found at the basal body. Scale bar = 10  $\mu$ m. Inset box shows the indicated region at higher magnification.

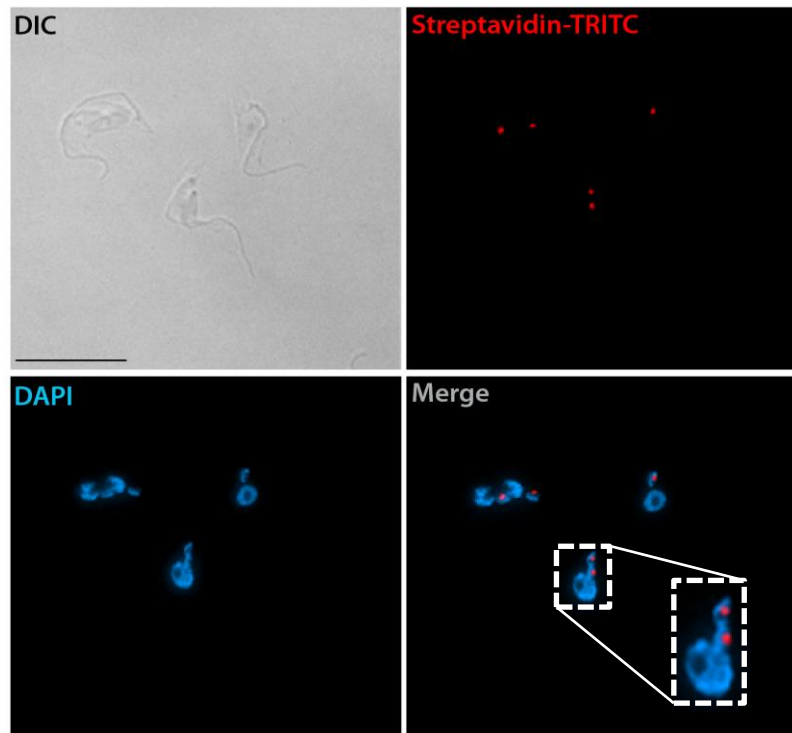


**Figure 3.8 Immunofluorescence analysis of a detergent extracted cells expressing *TbRP2::myc::BirA\**.** Anti-myc antibodies were used to detect the *TbRP2::myc::BirA\** protein. No expression can be detected in non-induced cells. In induced cells, fusion protein was clearly observed in the region of the basal body. Scale bars = 10  $\mu$ m. Inset box shows the indicated region at higher magnification.

Uninduced

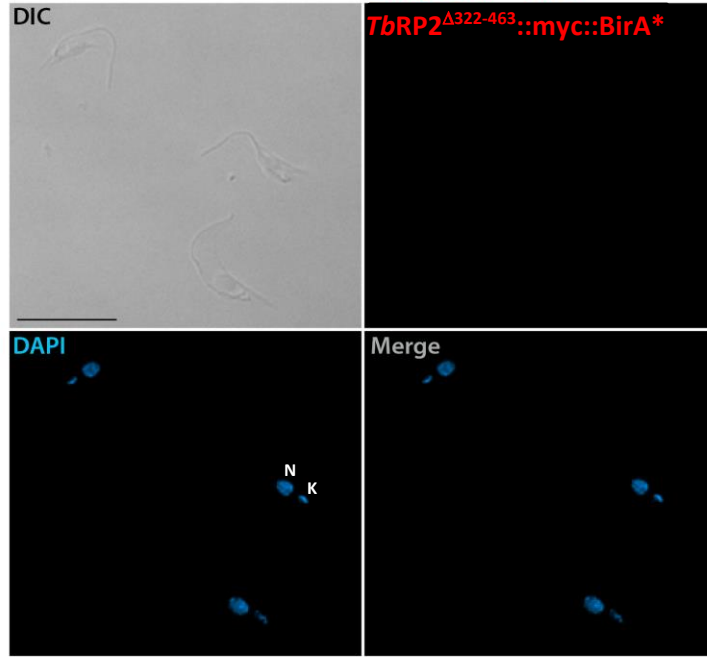


Induced

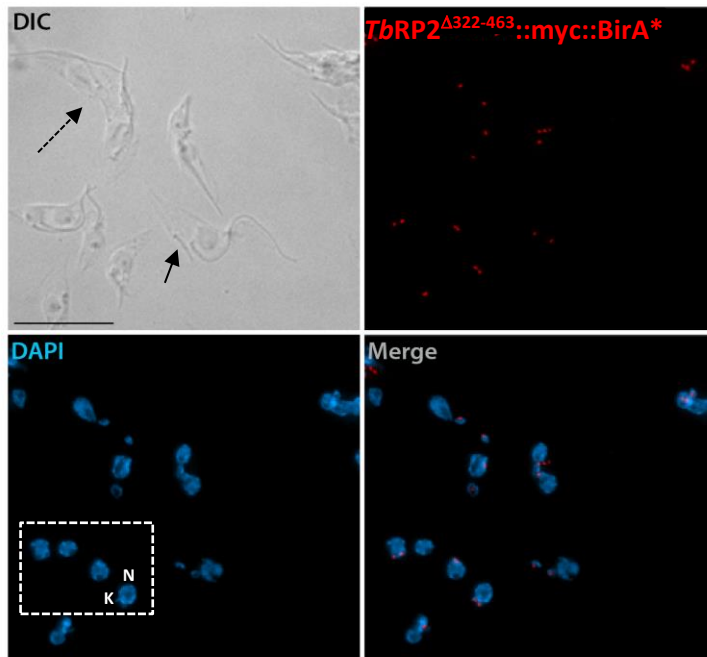


**Figure 3.9 Immunofluorescence analysis of detergent extracted cells expressing *TbRP2::myc::BirA\** probed with TRITC-conjugated streptavidin.** Weak nuclear staining can be observed in non-induced cells. However, after induction, biotinylated proteins can only be found at the basal body. Scale bar = 10  $\mu$ m. Inset box shows the indicated region at higher magnification.

Uninduced

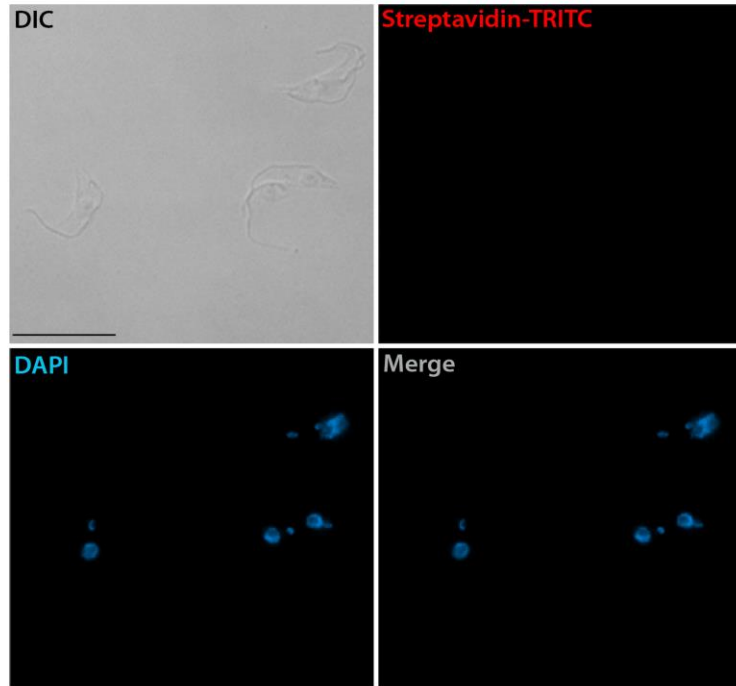


Induced

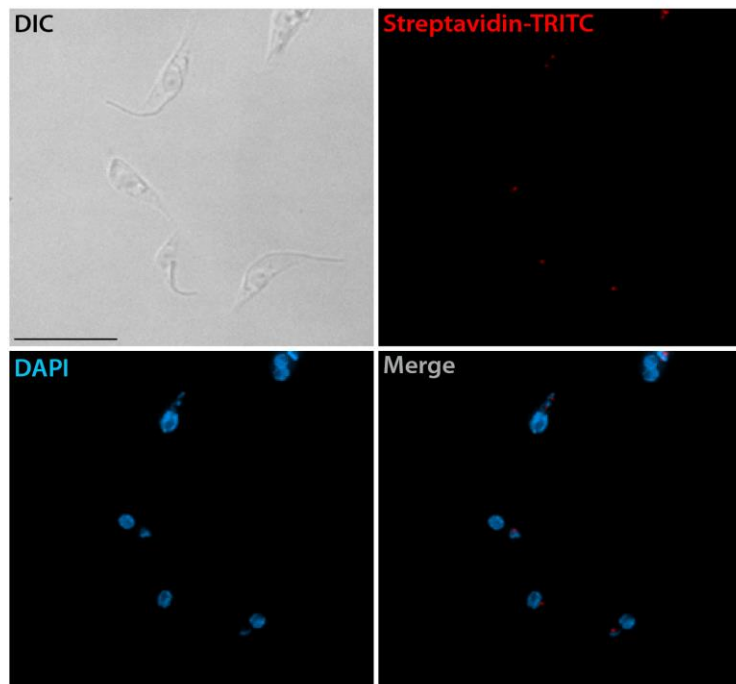


**Figure 3.10 Immunofluorescence analysis of a cell inducible expressing *TbRP2*<sup>Δ322-463</sup>::myc::BirA\* in the presence of excess biotin in detergent extracted fashion. Anti-myc antibodies were used to detect the tagged *TbRP2*<sup>Δ322-463</sup>::myc::BirA\*. No expression can be detected in non-induced cell. In induced cells, fusion protein was clearly observed in the region of basal body. Scale bars = 10 μm. Kinetoplast (K), Nucleus (N), black arrow solid line indicates short new flagellum and black arrow broken line indicates division defect. White broken-line rectangular shows cells that K is very close to N.**

Uninduced

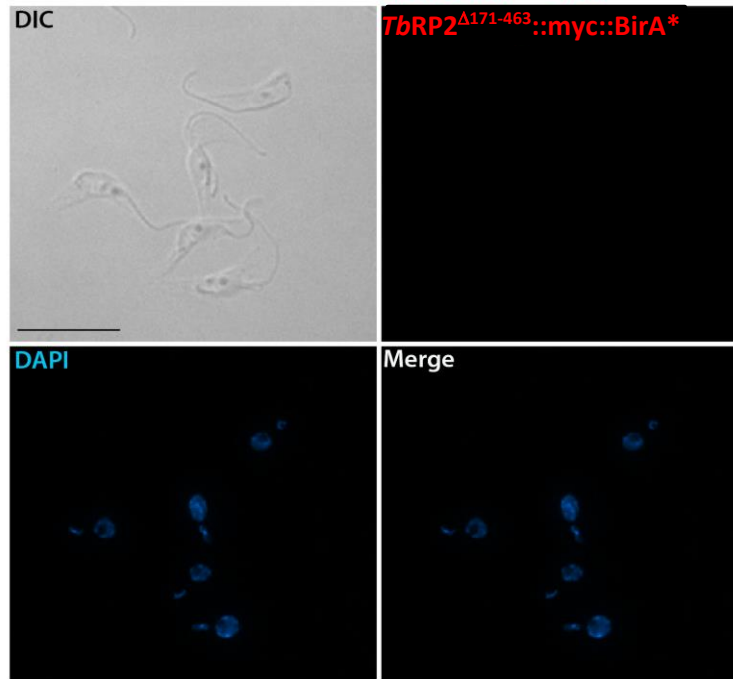


Induced

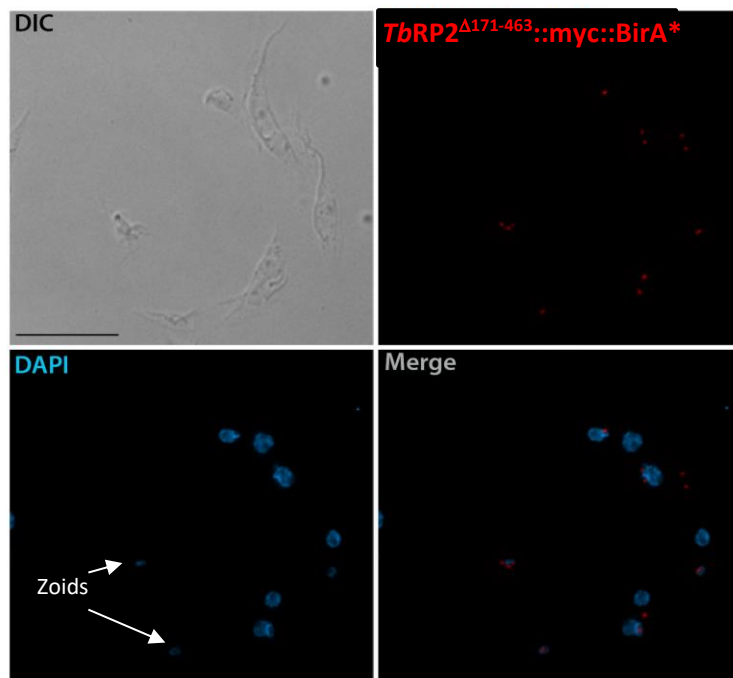


**Figure 3.11** Immunofluorescence analysis of detergent extracted cells expressing *TbRP2*<sup>Δ322-463::myc::BirA\*</sup> probed with TRITC-conjugated streptavidin. Weak nuclear staining can be observed in non-induced cells. However, after induction, biotinylated proteins can only be found at the basal body. Scale bar = 10 μm.

Uninduced



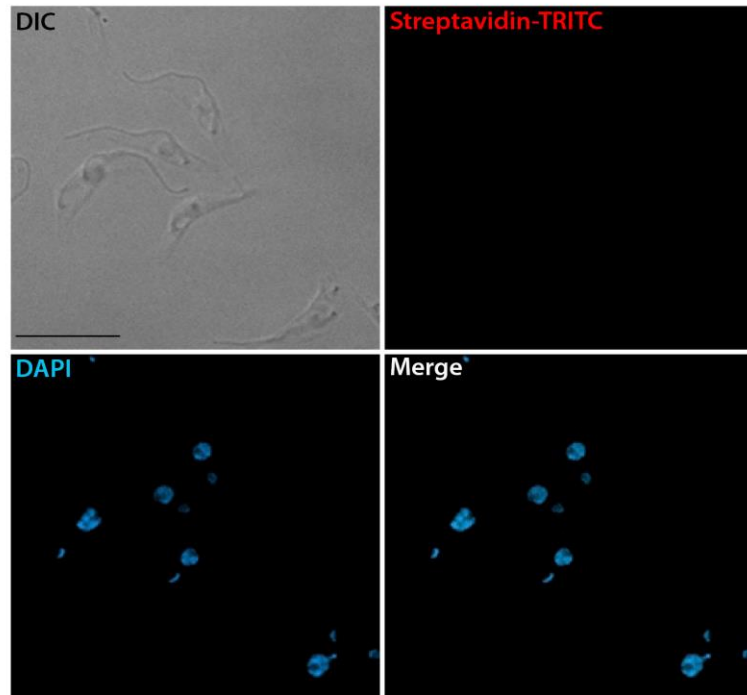
Induced



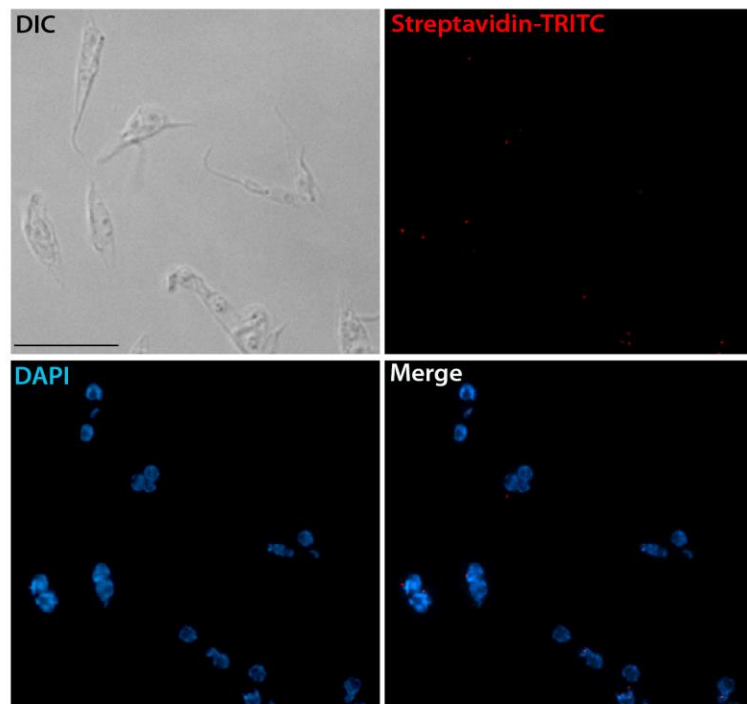
**Figure 3.12 Immunofluorescence analysis of a detergent extracted cells expressing *TbRP2*<sup>Δ171-463</sup>::myc::BirA\*.** Anti-myc antibodies were used to detect the *TbRP2*<sup>Δ171-463</sup>::myc::BirA\* protein. No expression can be detected in non-induced cells. In induced cells, fusion protein was clearly observed in the region of the basal body. Scale bars = 10  $\mu$ m.



Uninduced



Induced

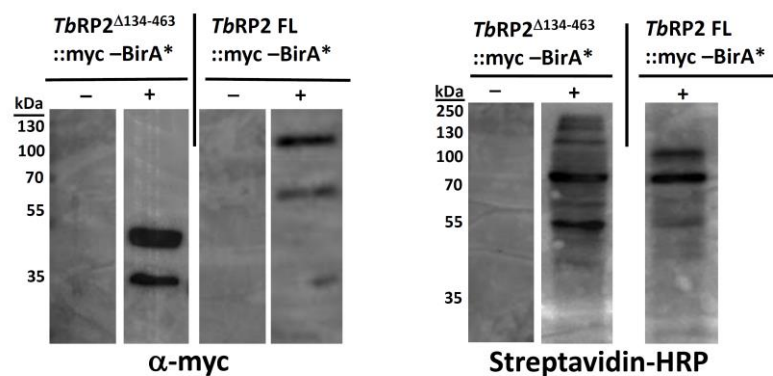


**Figure 3.13.** Immunofluorescence analysis of detergent extracted cells expressing *TbRP2*<sup>Δ171-463::myc::BirA\*</sup> probed with TRITC-conjugated streptavidin. Weak nuclear staining can be observed in non-induced cells. However, after induction, biotinylated proteins can only be found at the basal body. Scale bar = 10 μm.

### 3.5 Analysis of expressing *TbRP2::myc::BirA\** and *TbRP2*<sup>Δ134-463</sup>::*myc::BirA\** in *T. brucei* by immunoblotting

To determine whether the biotinylated protein profiles in *TbRP2* FL::*myc::BirA\** and *TbRP2*<sup>Δ134-463</sup>::*myc::BirA\** cells were similar or different, samples derived from non-induced and induced cells were resolved by SDS-PAGE and immunoblotted using streptavidin-HRP and anti-myc antibody (Fig 3.14). When samples were probed with anti-myc antibody, no protein expression was observed in the absence of doxycycline. Following induction, the expected band of *TbRP2*<sup>Δ134-463</sup>::*myc::BirA\** (~50kDa) was detected. However, a clear band around 35kDa was also detected, which most likely represents a C-terminally tagged proteolytic cleavage product of myc-tag. A band of the expected size of *TbRP2* FL::*myc::BirA\** (85kDa) was also detected by the anti-myc antibody. Two additional bands were also detected (~70kDa and 35kDa respectively-products that also are likely to represent processed cleavage products of myc tag. However, immunofluorescence experiments confirmed that both FL and truncated *TbRP2::myc::BirA\** only localised to the basal body and so were suitable for further BiolD experiments.

Biotinylated proteins derived from both FL and *TbRP2*<sup>Δ134-463</sup> were analysed by probing replicate immunoblots with HRP-conjugated streptavidin. This confirmed that induction of full length *TbRP2::myc::BirA\** and *TbRP2*<sup>Δ134-463</sup>::*myc::BirA\** resulted in a complexity of biotinylated proteins in these cells. No biotinylated proteins were detected in absence of doxycycline.

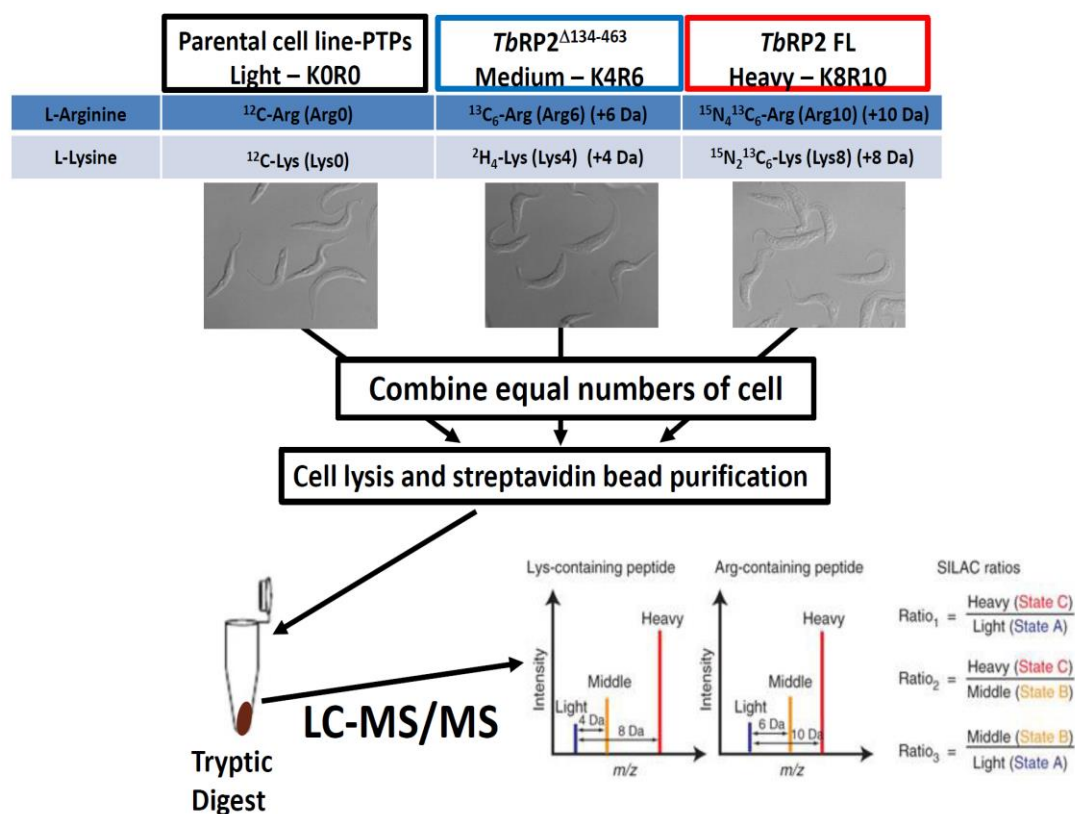


**Fig 3.14 Expression Immunoblots of *TbRP2* $^{\Delta 134-463}$ ::myc::BirA\* and *TbRP2* FL::myc::BirA\*.** Panel on the left is showing blots probed with anti-myc antibody. No expression can be observed in non-induced (-) cell populations. Following induction (+) bands of the the expected size of *TbRP2* $^{\Delta 134-463}$ ::myc::BirA\* (~ 49kDa) and *TbRP2* FL::myc::BirA\* (~85kDa) can be seen, in addition to some cleavage products. Biotinylated proteins were detected using HRP-conjugated streptavidin (Panel on the right). Following induction, a complex profile of biotinylated proteins was observed.

### 3.6 Stable Isotope Labelling with Amino acid in the Cell culture (SILAC)

Mass spectrometry-based proteomics has been used as a powerful tool for identifying and quantifying the components of multi-protein complexes for many years. However, errors can be introduced during protein fractionation or purification steps. To provide greater confidence in proteins identified in my BioID experiments, I used a quantitative proteomics method called stable isotope labelling of amino acid in cell culture (SILAC). This involves a metabolic labelling method that utilises the biological incorporation of stable isotopic labels into living cells (Ong et al., 2002). Although SILAC was originally described for mammalian cell culture (Ong et al., 2002), it has been adapted to perform in other different organisms such as yeast (Gruhler et al., 2005; Jiang and English, 2002), bacteria (Kerner et al., 2005), plants (Gruhler et al., 2005), *Plasmodium falciparum* (Nirmalan et al., 2004) and *T. brucei* (Urbaniak et al., 2012). Basically, cells are grown in medium which contains the amino acids arginine and/or lysine labelled with either normal, medium or heavy isotopes of  $^{13}\text{C}$  and/or  $^{15}\text{N}$ . The SILAC method relies on the complete incorporation of heavy amino acids during protein turnover (Ong and Mann, 2005, 2006).

The SILAC approach has several advantages; such as different cell populations can be mixed and treated as a single sample, which allows for downstream sample preparation without the risk of introducing quantification errors. Since the cells are labelled with different amino acids (and C or N isotopes) this will create a distinct and predictable mass difference between proteins derived from different samples. During mass spectrometry analysis, the relative abundance of proteins from different samples can be calculated from the intensities of 'heavy, medium and light' peptides. SILAC can discriminate potential contaminants purified by the BioID approach from true *TbRP2*-interacting and near neighbour proteins, which gives greater confidence in the identified proteins (see Fig 3.15 for the SILAC workflow).

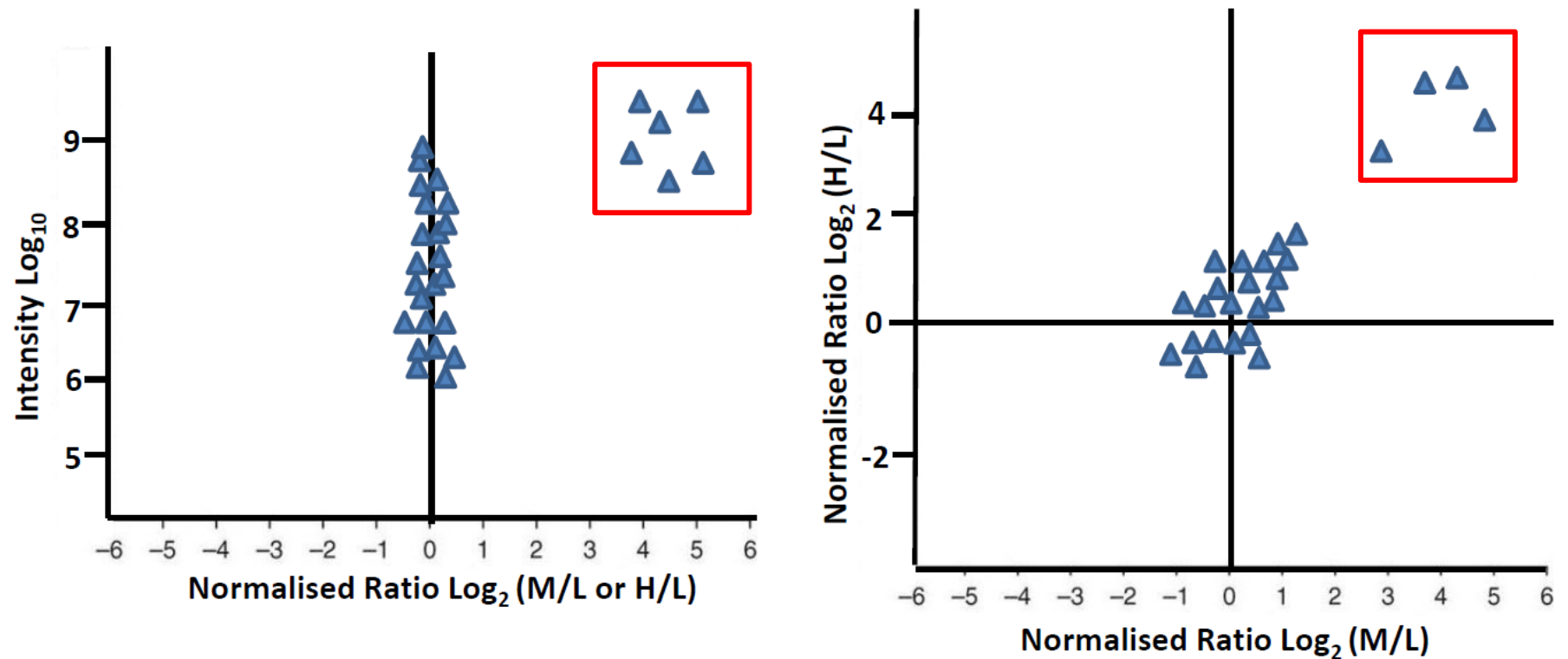


**Figure 3.15 SILAC workflow.** Three *T. brucei* cell lines were metabolically labelled with either (1) **light label (KOR0)** containing <sup>12</sup>C-Lysine and <sup>12</sup>C-Arginine, (2) **medium label (K4R6)** containing <sup>13</sup>C-Arginine and 4D-Lysine (3) **heavy label (K8R10)** containing <sup>13</sup>C/<sup>15</sup>N-Arginine and <sup>13</sup>C/<sup>15</sup>N-Lysine respectively for at least seven cell doublings prior to harvest. Further procedures were the same as described earlier.

### 3.7 Validation of TbRP2-interacting proteins and near neighbours by BioID and SILAC

Several preliminary experiments were performed to optimise the experimental conditions for analysis. Protein purification procedures were as described in Chapter 2 - and digested peptide samples were analysed at the FingerPrints Proteomic Facility at the University of Dundee by liquid chromatography tandem mass spectrometry (LC-MC/MS). The raw data obtained was processed using MaxQuant software (Cox and Mann, 2008) and the Perseus framework was used to complete proteomics analysis.

In theory, for any proteins identified by this mass spectrometry analysis, the relative intensity of abundance of each protein could be calculated. The abundance of each protein was determined by the total of all individual medium and light, or heavy and light peptide intensities detected for each protein (on the *y-axis* using a  $\log_{10}$  scale). The relative fold enrichment is shown using the  $\log_2$  ratio (medium/light or heavy/light normalised ratio to avoid any protein loading errors) on the *x-axis* (e.g. Fig 3.16 left). By using SILAC in combination with BioID it is possible to distinguish background noise (i.e. contaminants) from putative interacting and nearby partner proteins. The SILAC ratio of non-specific proteins/contaminants should be 1:1; i.e. proteins would lie around the 0 line in the figure. In contrast, 'real' TbRP2-interacting partner proteins would appear at the far right of the 0 line (Fig 3.16 left – red circle). When comparing 'medium-labelled versus light-labelled' against 'heavy-labelled versus light-labelled' (Fig 3.16 right), proteins which locate near the zero line from the *y-axis* are most likely an anomalous result. Those greatly enriched proteins from the experiment are more likely position at the far right hand corner (Fig 3.16 right – red circle).



**Figure 3.16** Example of a LC-MS/MS scatter plot showing the proteins identified in the SILAC-enriched ‘medium or heavy’ cell line ratio versus proteins identified in the ‘light’-labelled parental cell line plotted with respect to intensity (left). On the right, The LC-MS/MS scatter plot comparing proteins identified in the SILAC-enriched ‘medium’-labelled cell line ratio versus ‘light’-labelled parental cell line against proteins identified in ‘heavy’-labelled cell line ratio versus ‘light’-labelled parental cell line. Proteins, which locate around 0 line are more likely to be non-specific or contaminants. Proteins in the red square are highly enriched and are likely to be ‘true’ putative interacting or nearby proteins.

		Parental label-PTP	<i>TbRP2</i> <sup>Δ134-463</sup> Medium label	<i>TbRP2</i> FL Heavy label	<i>TbRP2</i> FL Medium label	Digestion
<b>WC-1</b>		R0K0	R6K4	R10K8		In Gel
<b>Cyto-1</b>	A	R0K0	R6K4	R10K8		On Beads
	B	R0K0	R6K4	R10K8		On Beads
<b>Cyto-2</b>	A	R0K0	R6K4	R10K8		On Beads
	B	R0K0	R6K4	R10K8		On Beads
<b>WC-2</b>	A	R0K0		R6K4	R10K8	On Beads
	B	R0K0		R6K4	R10K8	On Beads

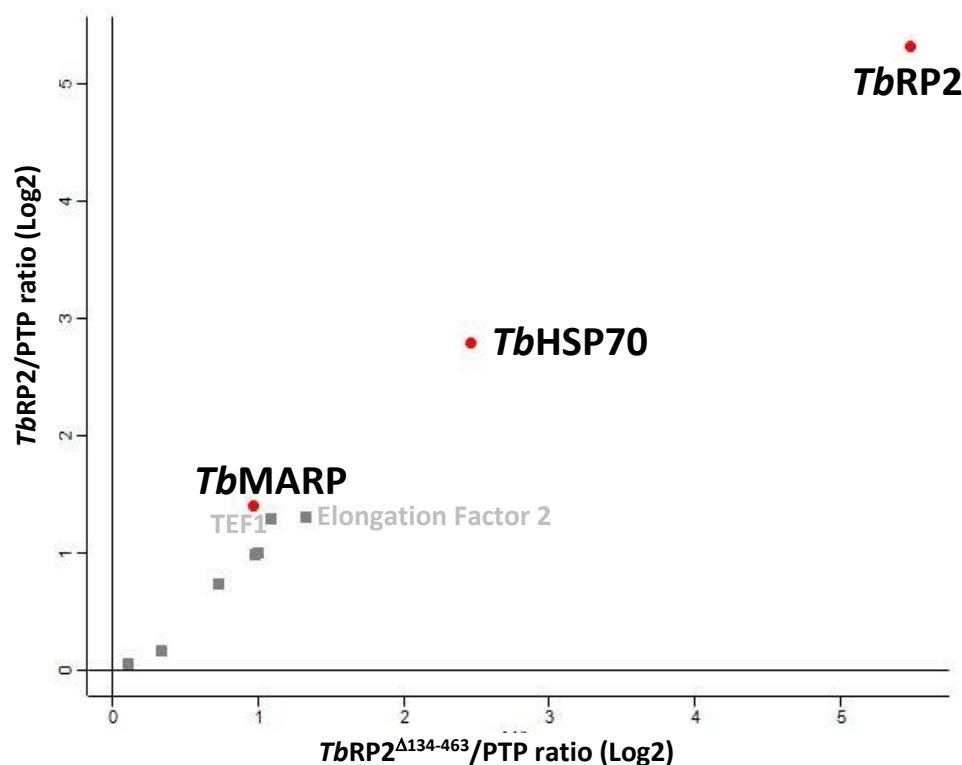
**Table 3.1 Summary of SILAC experiments performed in this thesis.** Expt. WC-1 represents the first experiment performed by using whole cell samples. Expt. Cyto-1/2 represents two biological replicates using cytoskeleton samples - each one also had one technical replicate performed on the sample. Expt. WC-2 represents the last experiment performed in this study using whole cell samples, but was undertaken only with *TbRP2* FL fusion protein as a label-swap experiment. *T. brucei* procyclic forms were labelled with unlabelled L-Arg and L-Lys, designated as R0K0, or with [<sup>13</sup>C<sub>6</sub>]-L-Arg and [<sup>2</sup>H<sub>4</sub>]-L-Lys, designated as R6K4, or with [<sup>15</sup>N<sub>4</sub><sup>13</sup>C<sub>6</sub>]-L-Arg and [<sup>15</sup>N<sub>2</sub><sup>13</sup>C<sub>6</sub>]-L-Lys, designated as R10K8. Tryptic peptides were generated either by trypsin in-gel digestion of protein bands from SDS-PAGE or in solution (on bead digestion) for analysis by MS.

### 3.7.1 Experiment 1: Whole cell -1, in gel digestion

The first BioID-SILAC experiment was performed with whole cell samples using in gel digestion to prepare peptides for MS analysis. The PTP parental cell line was grown in SDM-79-SILAC containing light label ((arginine (R0) and lysine (K0)). The *TbRP2*<sup>Δ134-463</sup>::myc::BirA\* cell line was grown in SDM-79-SILAC containing medium label (arginine (R6) and lysine (K4)), and *TbRP2* FL::myc::BirA\* cell line in media containing heavy label (arginine (R10) and lysine (K8)). Unfortunately, few proteins with a significant SILAC ratio were identified in this experiment (Fig 3.17). However, the *TbRP2* protein was the top protein in the ID list (See appendix Excel files, CD), and in the graph in Fig 3.17



*TbRP2* is positioned at top right corner (i.e. most enriched), which was a good indication that the BioID approach had worked. There were some interesting proteins, such as  $\alpha\beta$ -tubulin, various chaperones and a large highly repetitive protein annotated in the *T. brucei* genome as “microtubule-associated proteins” (Table 3.2). This microtubule-associated protein was particular interesting as it could potentially be a component of the transitional fibres and was selected for localisation (Chapter 5) and functional studies (Chapter 6).



**Figure 3.17** LC-MS/MS scatter plot comparing proteins identified in the SILAC-enriched ‘Medium’-labelled *TbRP2* $\Delta$ 134-463::myc::BirA\* cell line ratio versus ‘Light’-labelled PTPs parental cell line on the x-axis, against proteins identified in ‘Hheavy’-labelled *TbRP2*FL::myc::BirA\* cell line ratio versus ‘Light’-labelled PTPs parental cell line on the y-axis. The samples were prepared in whole cells. The peptide(s) derived from each protein (point on the plot graph) were detected by mass spectrometry and scales are logged so data is normally distributed and the ratios normalised to avoid any loading errors. *TbRP2* shows the most SILAC enrichment (i.e. top right corner).

Accession number	Product	Ratio M/L	Ratio H/L	Unique peptides	Sequence coverage [%]	Unique sequence coverage [%]	Mol. weight [kDa]	Intensity	Intensity L	Intensity M	Intensity H
Tb927.10.14010	tubulin cofactor C domain-containing protein RP2 (RP2)	5.484	5.317	3	8.4	8.4	50.631	6.159	4.900	5.3911	6.048
Tb927.11.11330 ;Tb11.v5.1035	heat shock protein 70	2.466	2.790	10	21.3	21.3	71.408	7.117	6.455	6.5625	6.818
Tb927.10.4570 ;Tb927.10.4560	elongation factor 2	1.331	1.299	2	3.5	3.5	94.333	6.002	5.462	5.5695	5.537
Tb927.10.2110 ;Tb927.10.2100 ;Tb11.v5.1046 ;Tb927.10.2090	elongation factor 1-alpha (TEF1)	1.083	1.284	5	13.8	13.8	49.105	6.898	6.540	6.2199	6.443
Tb927.1.2390 ;Tb927.1.2370 ;Tb927.1.2350 ;Tb927.1.2330 ;Tb11.v5.0469 ;Tb927.1.2410	beta tubulin	0.998	0.993	16	51.8	51.8	49.704	7.809	7.366	7.2071	7.401
Tb927.1.2400 ;Tb927.1.2380 ;Tb927.1.2360 ;Tb927.1.2340	alpha tubulin	0.983	0.981	13	45	45	49.787	7.592	7.191	6.9507	7.165
Tb927.10.10360	microtubule-associated protein, putative	0.967	1.401	3	40.7	40.7	267.6	6.593	6.113	5.9514	6.237
Tb927.10.10980 ;Tb927.10.10970 ;Tb927.10.10960 ;Tb927.10.10950 ;Tb927.10.10940 ;Tb927.10.10930 ;Tb927.10.10920 ;Tb927.10.10910 ;Tb927.10.10900 ;Tb927.10.10890 ;Tb11.v5.0543	heat shock protein 90, putative	0.723	0.728	2	3	3	80.793	6.164	5.878	5.5222	5.569
Tb927.10.6690	protein tyrosine phosphatase, putative (TbPTP1)	0.103	0.053	6	27.8	27.8	34.132	7.148	7.125	5.6809	5.392

**Table 3.2** Proteins identified in *TbRP2* BirA\* fusion protein expressing *T. brucei* by BioID-SILAC (WC-1). Yellow highlight indicated bait protein *TbRP2*, and MARP protein in blue highlight was selected for further analysis. M/L: *TbRP2*<sup>Δ134-463</sup>/PTP; H/L : *TbRP2* FL /PTP

### 3.7.2 Experiment 2 - cytoskeletons, in solution/on bead digestion

In-gel digestion is considered to be a very reproducible and effective method for sample preparation, as it ensures that detergents and salts can be removed from samples and thus avoids any harm to the mass spectrometer (Rosenfeld et al., 1992). However, because in-gel digestion is very time-consuming (Park and Russell, 2001), and there can be loss of peptides during extraction steps (i.e. peptides remain trapped within the gel matrix) this approach can be problematic (Rosenfeld et al., 1992). Therefore, I decided to alter experimental protocols for all subsequent experiments and digest proteins with trypsin while they were still attached to the streptavidin beads.

In this experiment, SILAC labelling remained the same i.e. parental PTP cells -Light (R0K0); *TbRP2*<sup>Δ134-463</sup>::myc::BirA\*-Medium (R6K4) and *TbRP2* FL::myc::BirA\*-Heavy (R10K8). However, cultures were harvested using PEME 1% NP-40 extraction buffer containing protease inhibitors in the extraction buffer (see Materials and Methods for details) to prepare insoluble cytoskeleton samples (*TbRP2* is a detergent insoluble protein). This experiment was repeated twice and each time the sample was analysed twice (i.e. two biological and two technical replicates). The data from each individual experiment (Expt. cyto-1 A and B; Expt. cyto-2 A and B) and the combined data for all four experiments were analysed using MaxQuant software (Cox and Mann, 2008) and the EupathDB *T. brucei* protein database (Aslett et al., 2010).

For proteins identified in cyto-1 and 2 and the combined data set, a pairwise comparison was performed of the log<sub>2</sub> (M/L or H/L) ratios for proteins identified in the four individual or combined data sets. The Pearson's correlation coefficient values were calculated using Perseus 1.3.0.4 ([www.perseus-framework.org](http://www.perseus-framework.org)) (see appendix I Fig 9.1-9.3). According to the Pearson coefficient analysis, there is a good correlation between *TbRP2*<sup>Δ134-463</sup>/PTP and *TbRP2* FL/PTP with Expt. cyto-1 A and B (Pearson coefficients above 0.8). However, technical replicates (between A and B) have weaker

correlations (Pearson coefficients around 0.5) (Appendix I Fig 9.1). This could be due to the fact that when only a few proteins have been identified by MS, Pearson coefficients only take the enriched proteins into account (i.e. there was not enough data to provide a strong correlation and so the scatter plots looked worse). Again, similar results were obtained to the Expt. cyto-2 A and B; there is very good correlation between *TbRP2*<sup>Δ134-463</sup>/PTP and *TbRP2* FL/PTP in both Expt. cyto-2 A and B (Pearson coefficients above 0.9). In this second set of technical replicates (between cyto-2 A and B), the Pearson coefficient is around 0.7 (Appendix I Fig 9.2). When I combined the Expt. cyto-1 A and B together (*TbRP2*<sup>Δ134-463</sup>/PTP from A and B; and *TbRP2* FL/PTP from A and B) to compare with Expt. cyto-2 A and B together (as biological replicates comparison), the Pearson coefficient analysis revealed that within cyto-1 and cyto-2, there was a relatively good correlation between *TbRP2*<sup>Δ134-463</sup>/PTP and *TbRP2* FL/PTP (~0.6-0.7). However, the correlation of biological replicates of *TbRP2*<sup>Δ134-463</sup>/PTP was less strong (~0.3), but *TbRP2* FL/PTP showed a better correlation value (~0.7) (Appendix I Fig 9.3).

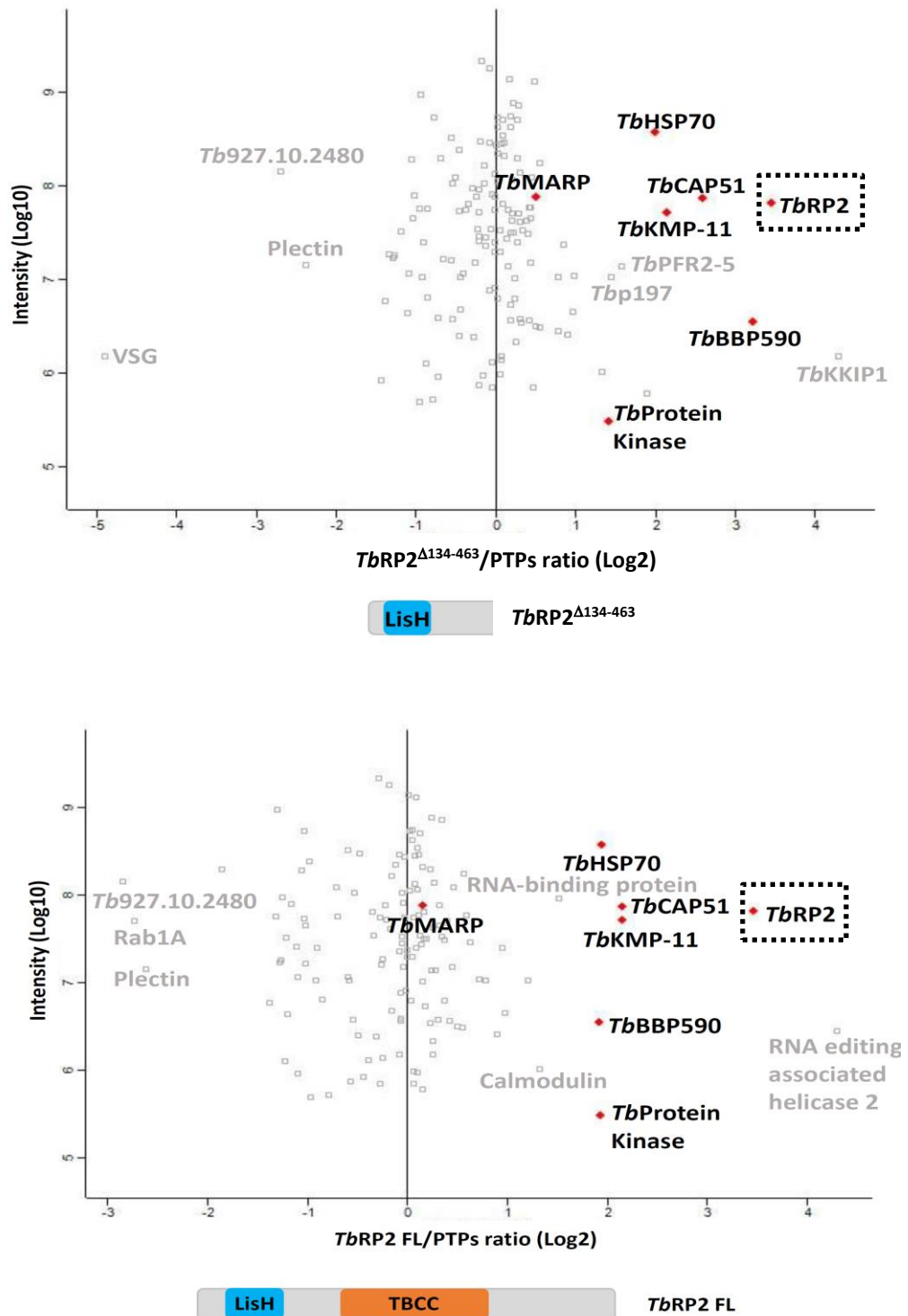
The proteins identified in the cyto 1+2 experiments (combined) is shown in Table 3.3. To prioritise appropriate proteins for further analysis, several initial bioinformatics approaches were used. Firstly, all proteins with a SILAC ratio above 1 were interrogated at GENEDB and TriTrypDB to determine what was known about the proteins, such as GO function(s) and occurrence in other proteomic datasets. Proteins with potential processing or structural roles were prioritised for further study (highlighted in blue in Table 3.3). As mentioned earlier, two types of scatter plot can be used to visualise the enrichment of proteins from the MS experiments (Fig 3.16). When comparing the *TbRP2*<sup>Δ134-463</sup>::myc::BirA\* cell line (medium-label) versus the PTPs cell line (light-label) (Fig 3.18-Top), proteins were considered to be enriched when they were above the cut-off threshold (SILAC ratio above 1). This included *TbRP2* itself (red diamonds, black square), which again was a good indication that the experiment had worked. Other proteins (red diamonds) were above the threshold, and so were selected for further study; these included *TbCAP51*, *TbKMP-11*, *TbBBP590*, *TbHSP70* and *TbProtein Kinase* (see bioinformatics analysis in Chapter 4 for

further rationale as to why these proteins were selected for study). The *Tb*MARP protein, which was identified previously in Expt1 WC-1 (Fig 3.17), is positioned much closer to the zero line (i.e. background noise) in this scatter plot. Thus, *Tb*MARP would not have been selected for further study on the basis on this cytoskeleton experiment. However, by the time I carried out and analysed Expt. cyto-1 -2 data, DNA constructs, and trypanosome cell lines had already been generated and so the decision was taken to continue to investigate the *Tb*MARP protein (see Chapters 5 and 6). The scatter plot of each protein on the graph also reflected the heavy-labelled *Tb*RP2 FL::myc::BirA\* versus light-labelled PTPs (Fig 3.18-bottom). Here, the selected proteins are also marked as a red diamond-shape on the graph. *Tb*RP2 was one of the most enriched proteins located above the cut off threshold, and again *Tb*MARP was in the region categorised as noise. In Fig 3.19, the plot shows the comparison between the 'medium-labelled versus light-labelled' against 'heavy-labelled versus light-labelled' (Fig 3.19), all of the proteins selected for further study were highly enriched (i.e. located at the top right-hand corner of the graph).

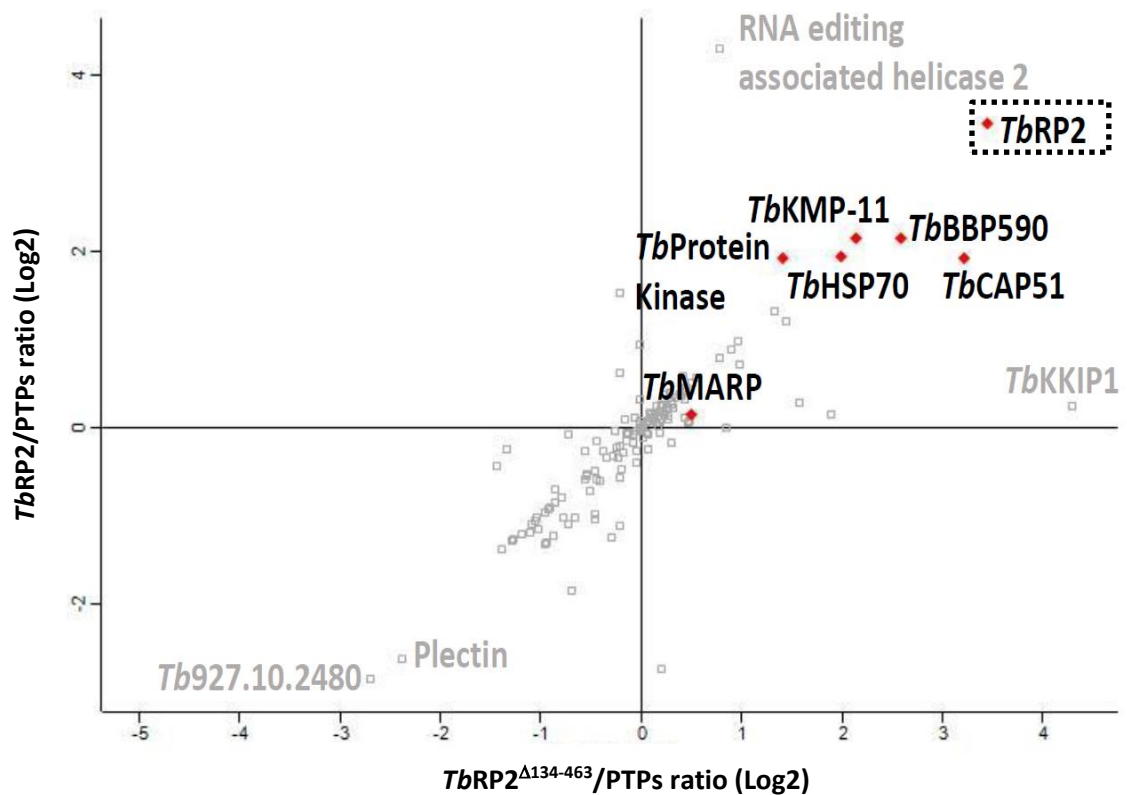
So far I have just presented scatter plots to give an overview of the proteins identified; i.e. where they are positioned on the plots based on their enrichment. In the following graphs, I illustrate how reproducible (based on SILAC ratios) the identification of the selected proteins were in the two biological experiments (cyto-1 and 2) (Fig 3.20). The SILAC ratios ( $\log_2$ ) on the y-axis show the combined value from all four experiments (cyto-1A; B and cyto-2A; B). The graphs show that the SILAC ratios for *Tb*RP2 $\Delta^{134-463}$ /PTP in blue and *Tb*RP2 FL/PTP in dark grey are very similar between the two biological experiments. However, there are some small differences between the SILAC ratios for *Tb*HSP70 and *Tb*Protein Kinase, and *Tb*KMP11 and *Tb*MARP show larger differences between the two biological experiments. For *Tb*CAP51, the M/L SILAC ratios are very similar between two biological experiments, but not for the H/L labelling experiment. In addition, there are big differences between M/L and H/L in cyto-1 and 2, and *Tb*BBP590 appears to have the same pattern as *Tb*CAP51, with proteins more enriched in the *Tb*RP2 $\Delta^{134-463}$  experiment.

Majority protein IDs	Products	M/L normalized	H/L normalized	Intensity	Intensity L	Intensity M	Intensity H	Mol. weight [kDa]	Sequence coverage [%]	Peptide counts (all)	Peptide counts (unique)
Tb927.4.1500	RNA editing associated helicase 2 (REH2)	0.778	4.301	6.445	5.214	5.240	6.389	241.82	0.7	1	1
Tb927.10.14010	tubulin cofactor C domain-containing protein RP2 (RP2)	3.456	3.456	7.820	6.855	7.455	7.483	50.631	42.5	14	14
Tb927.7.2640;Tb927.7.2650	CAP51	2.589	2.150	7.877	7.597	7.235	7.268	50.853	16.3	8;4	8;4
Tb927.9.13920;Tb927.9.13880;Tb927.9.13820	kinetoplastid membrane protein KMP-11 (KMP-11)	2.145	2.145	7.719	7.286	7.214	7.224	11.076	60.9	6;6;6	6;6;6
Tb927.11.11330	heat shock protein 70	1.995	1.945	8.581	7.995	8.071	8.215	71.408	34.6	18	18
Tb927.9.6560	protein kinase, putative	1.403	1.928	5.485	4.653	4.992	5.211	80.198	1.8	1	1
Tb927.11.10660	Tb BBP590	3.212	1.920	6.546	5.568	6.312	6.040	81.219	5.9	1	1
Tb927.2.4710	RNA-binding protein, putative (TRRM1)	-0.209	1.517	7.963	7.695	7.279	7.368	49.983	25.4	8	8
Tb927.11.13050;Tb927.11.13040;Tb927.11.13030;Tb927.11.13020	calmodulin	1.323	1.323	6.015	5.409	5.591	5.591	16.838	20.1	2;2;2;2	2;2;2;2
Tb927.10.15750	p197, a TAC component	1.451	1.212	7.025	6.604	6.512	6.520	183.78	12.5	4;4	4;4
Tb927.5.930;Tb11.v5.0613;Tb927.5.940	NADH-dependent fumarate reductase (FRDg)	0.972	0.972	6.648	6.174	6.165	6.173	123.69	3.1	3;2;2	3;2;2
Tb927.3.4330;Tb927.3.4320;Tb927.3.4310;Tb927.3.4300;Tb927.3.4290;Tb11.v5.1055	73 kDa paraflagellar rod protein, PFR1 (PFR1)	-0.013	0.943	7.397	7.049	6.636	6.974	68.682	2.2	1;1;1;1;1;1	1;1;1;1;1;1
Tb927.10.15700	hypothetical protein, conserved	0.890	0.890	6.410	5.965	5.916	5.916	377.23	9	1	1
Tb927.3.3270	ATP-dependent phosphofructokinase (PFK)	0.781	0.781	7.031	6.651	6.418	6.562	53.517	19.5	8	8
Tb11.v5.0410;Tb927.4.2070;Tb11.v5.0881	antigenic protein, conserved	0.982	0.721	7.036	6.651	6.499	6.510	200.51	47	4;4;3	4;4;3
Tb927.8.900	splicing factor TSR1 (TSR1)	-0.215	0.622	7.460	7.165	6.791	6.905	37.469	15.9	5	5
Tb927.11.7500;Tb927.11.7490	hypothetical protein, conserved	0.414	0.586	7.771	7.472	7.001	7.287	13.594	7.8	1;1	1;1
Tb927.7.5020;Tb927.7.5000;Tb11.v5.0182;Tb11.v5.0181	60S ribosomal protein L19, putative	0.546	0.557	8.246	7.959	7.627	7.631	29.408	14.2	4;4;4;4	4;4;4;4

Table 3.3 Proteins identified in *TbRP2* BirA\* fusion protein expressing *T. brucei* by BioID-SILAC (Cyto 1+2). Yellow highlight indicated bait protein -*TbRP2*, and proteins highlighted in blue were selected for further analysis. M/L : *TbRP2*<sup>Δ134-463</sup>/PTP; H/L : *TbRP2* FL /PTP.



**Figure 3.18** Scatter plot showing proteins identified in the SILAC-enriched ‘medium’-labelled *TbRP2*<sup>Δ134-463</sup>-BirA\* cell line (upper panel) or ‘heavy’-labelled *TbRP2* FL-BirA\* cell line (lower panel). The x-axis shows the log2 normalised ratio of ‘medium or heavy’ labelled proteins versus ‘light’-labelled proteins, and the y-axis represents the relative abundance of each protein.



**Figure 3.19** LC-MS/MS scatter plot comparing proteins identified in the SILAC-enriched 'medium'-labelled *TbRP2*<sup>Δ134-463</sup>::myc::BirA\* cell line ratio versus 'light'-labelled PTP parental cell line against proteins identified in 'heavy'-labelled *TbRP2* FL::myc::BirA\* cell line ratio versus 'light'-labelled PTPs parental cell line. The x-axis shows the ratio of 'medium'-labelled proteins versus 'light'-labelled proteins, identified according to the peptides detected by mass spectrometry. The y-axis shows the ratio of 'heavy'-labelled proteins versus 'light'-labelled proteins, identified according to the peptides detected by mass spectrometry. Scales is normally distributed and the ratios normalised because the base peak originally lay to the left of zero. Proteins, which appear in the centre are more likely to be contaminant.



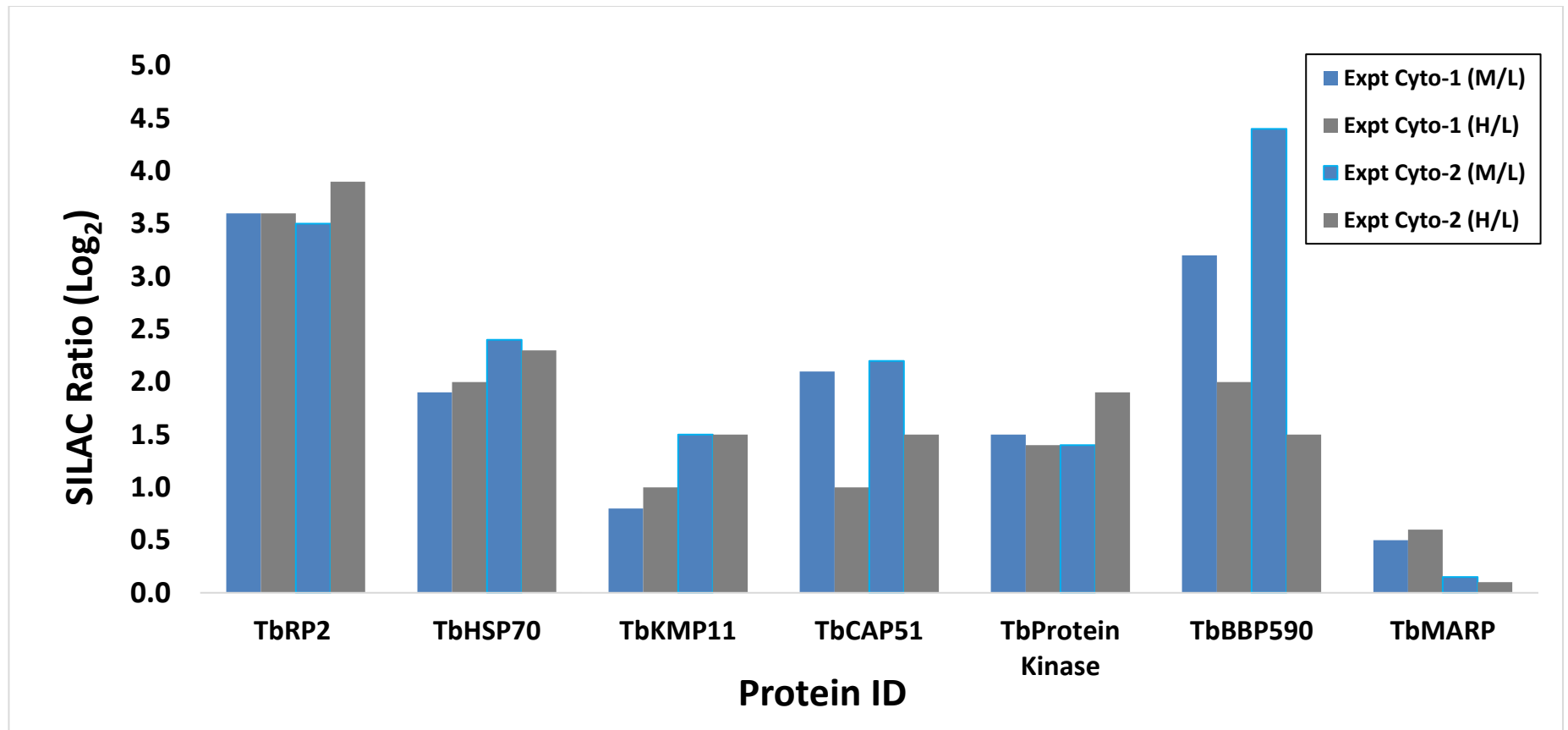


Figure 3.20 The SILAC ratio of each selected proteins in the BioID cytoskeleton cell fraction is indicated by the colour of each bar. M/L = *TbRP2*<sup>Δ134-463</sup>/PTP; H/L = *TbRP2*<sup>Δ134-463</sup>/PTP

### 3.7.3 Experiment 3- whole cell, in solution/on bead digestion

Based on the previous data, the number of the proteins identified using an on beads digestion protocol was much greater than in the first (in gel digestion) experiment and identified a number of proteins of interest that were selected for further study. However, one of my initial aims was to interrogate how *TbRP2* was targeted to the basal body solely by the N-terminus TOF-LisH motifs, and so I was interested to see whether different protein profiles were identified in cells expressing *TbRP2*<sup>Δ134-463</sup> or full length *TbRP2*. Based on the analyses I have just shown (i.e. cyto expts), little difference was observed between the peptides identified from medium and heavy labelled samples i.e. cells expressing *TbRP2*<sup>Δ134-463</sup> or full length *TbRP2*. Therefore, in subsequent experiments, I decided to use only cells expressing full length *TbRP2::myc::BirA\**; these cells were labelled using medium and heavy labels in a label swap experiment. Additionally, I decided to use whole cell preparations and not cytoskeletal extracts, as I was concerned that some *TbRP2*-interacting proteins/proximal proteins may be detergent soluble and would therefore be lost from my analysis.

The data from each individual experiment (WC-2 A, B) and the combined data for both experiments were analysed using MaxQuant software (Cox and Mann, 2008) and the EupathDB *T. brucei* protein database (Aslett et al., 2010). For the proteins identified in the both Expt WC-2 A, B and combined data set, I again performed a pairwise comparison of the log<sub>2</sub> (M/L or H/L) ratios with those for the proteins identified in the four individual or combined data sets. The Pearson's correlation coefficient values were calculated using Perseus 1.3.0.4 ([www.perseus-framework.org](http://www.perseus-framework.org)). This determined good correlations (Pearson coefficients are all above 0.9) between M/L and H/L of Expt. WC-2 A and B, indicating that the results are not biased (Appendix I Fig 9.4).

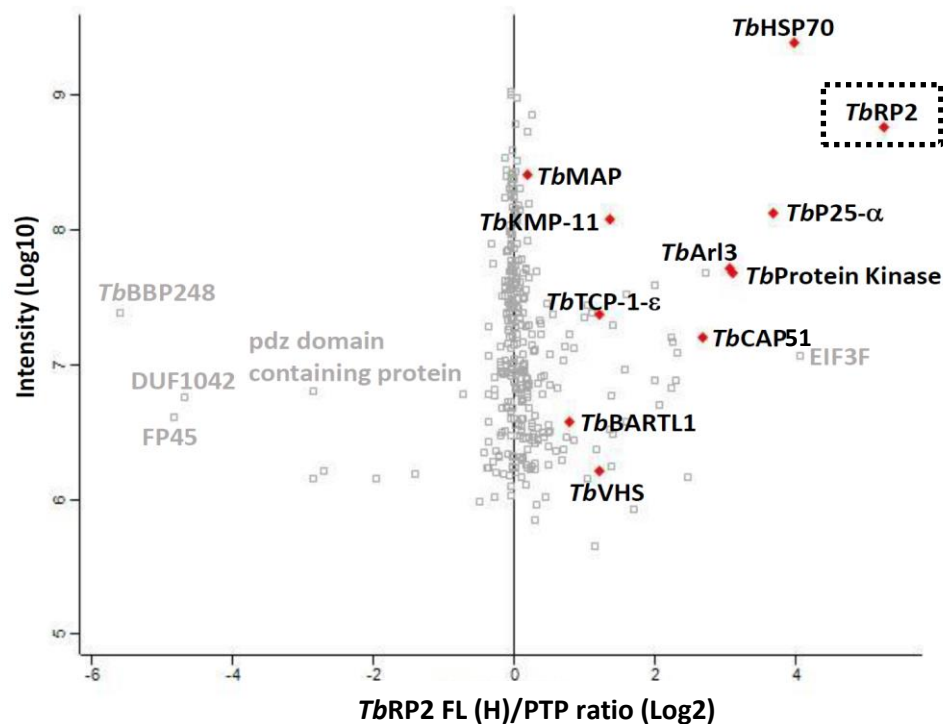
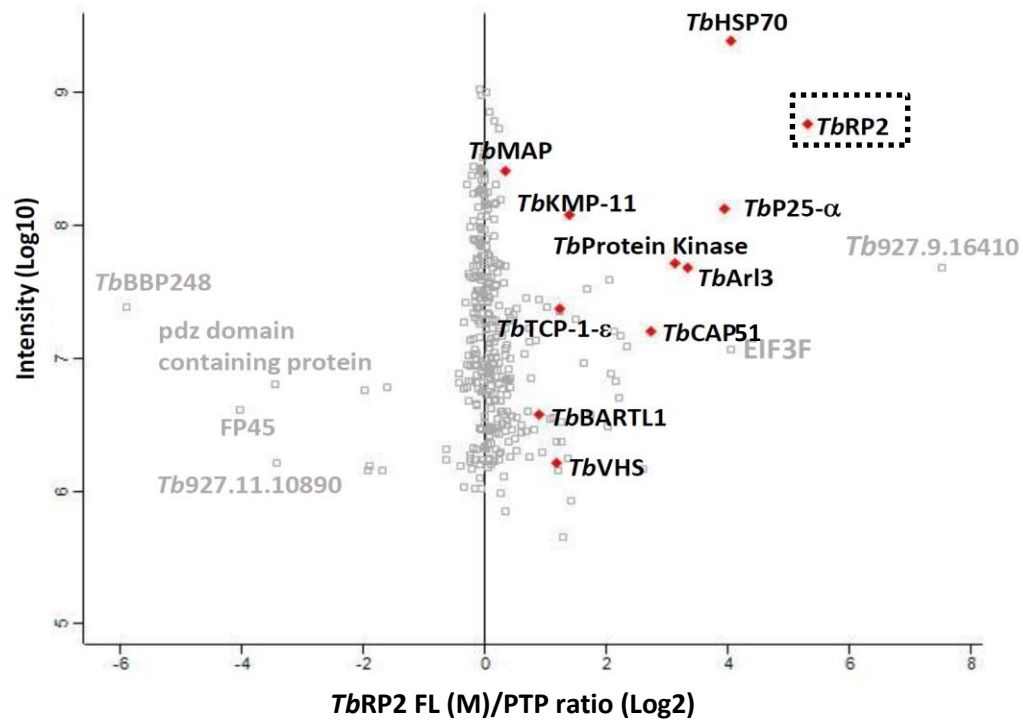
I compared the medium-labelled *TbRP2* FL-MYC-BirA\* cell line versus the light-labelled parental PTPs cell line (i.e. combined data; M/L: Expt-WC-A and B) against intensity (Log10) (Fig 3.21-Top). The number of proteins identified were considerably greater than when experiments were performed using cytoskeletal samples (Table 3.4). The *TbRP2* protein was again above the threshold showing protein enrichment, indicating the BioID/SILAC experiment had worked. Five out of six of the proteins selected from the previous experiments were also identified on the scatter plot (apart from *TbBBP590* protein), along with another five additional proteins that were selected for further study (labelled with red diamond-shaped on the plot). Again, the criteria were used in WC-2 was same as described in the previous Cyto experiment; except *BARTL1* (below SILAC ratio 1, see chapter 4.12), all other selected proteins with SILAC ratio  $\geq 1$ , bioinformatics analysis was also performed to prioritise candidates – again these were proteins with potential roles in protein processing. The proteins were *TbArl3*, *TbBartL1*, *TbP25- $\alpha$* , *TbTCP-1- $\epsilon$*  and *TbVHS-domain* containing protein (see Chapter 4 for further details). The scatter plot proteins identified in the ‘heavy-labelled’ *TbRP2* FL::myc::BirA\* sample (expressed as a ratio against ‘light-labelled’ PTPs is shown in Fig 3.21-bottom panel). Here, the selected proteins are also marked as a red diamond-shape on the graph and are remarkably similar in position to the scatter plot shown in Fig 3.22 top panel. This confirmed that the label had no effect on detection of the protein. *TbRP2* was one of the enriched proteins located above the threshold. In Fig 3.22, the plot shows the comparison between the ‘medium-labelled versus light-labelled’ against ‘heavy-labelled versus light-labelled’ (Fig 3.22), it is clear that all the selected proteins are greatly enriched (i.e. located at the right top corner). In Fig 3.23, I have combined all *TbRP2* FL Cyto data together (H/L) to compare with the WC data (H/L). *TbRP2* is still positioned at the right top corner (i.e. highly enriched), and other enriched proteins identified by both the Cyto and WC experiments are *TbHSP70*, *TbCAP51* and *TbMARF*. When I used the SILAC ratio generated from each experiment (Expt. WC-2 A and B, technical replicates) of each individual protein (Fig 3.24), the results are very similar between experiment, which suggested that Expt. WC 1 and 2 are reproducible.

Majority protein IDs	Products	M/L normalized	H/L normalized	Intensity	Intensity L	Intensity M	Intensity H	Mol. weight [kDa]	Peptide counts (all)	Peptide counts (unique)	Unique sequence coverage [%]
Tb927.10.14010	RP2	5.295	5.262	8.559	6.841	8.191	8.301	50.631	18	18	41
Tb927.3.1680	Eukaryotic translation initiation factor 3 subunit F (EIF3F)	4.206	4.285	6.803	5.336	6.404	6.556	35.191	1	1	2.8
Tb927.11.11330	heat shock protein 70	4.012	3.932	8.980	7.769	8.609	8.691	71.408	31	29	47
Tb927.4.2740	p25-alpha	3.768	3.460	7.787	6.749	7.391	7.493	16.327	10	10	53.3
Tb927.6.3650	ARL3 Like	3.120	3.054	7.431	6.401	7.068	7.106	20.072	5	5	30.1
Tb927.9.16410	hypothetical protein	7.934	2.950	7.801	5.293	7.785	6.315	12.544	1	1	14
Tb927.7.2640	CAP51	2.818	2.788	7.112	6.428	6.716	6.703	50.853	5	5	9.9
Tb927.11.6440	mRNA binding	2.125	2.306	6.463	5.648	6.009	6.157	48.841	2	2	6.6
Tb927.11.5840	Protein translation factor SUI1 homolog, putative	2.224	2.222	6.642	5.806	6.240	6.304	12.738	2	2	22.9
Tb927.9.5730 (previously known as : Tb09.160.4240)	nucleosome assembly protein-like protein	2.167	2.177	7.036	6.754	6.320	6.489	47.584	4	4	15.6
Tb927.9.6560	protein kinase, putative (pseudokinase family)	2.410	2.168	7.409	6.507	6.993	7.099	80.198	8	8	14.7
Tb927.8.2820	hypothetical protein, conserved	1.913	1.962	6.684	5.729	6.289	6.370	141.11	2	2	27.2
Tb11.02.5400	cystathionine beta-synthase, putative	2.059	1.820	6.989	6.094	6.603	6.653	39.464	5	5	25.7

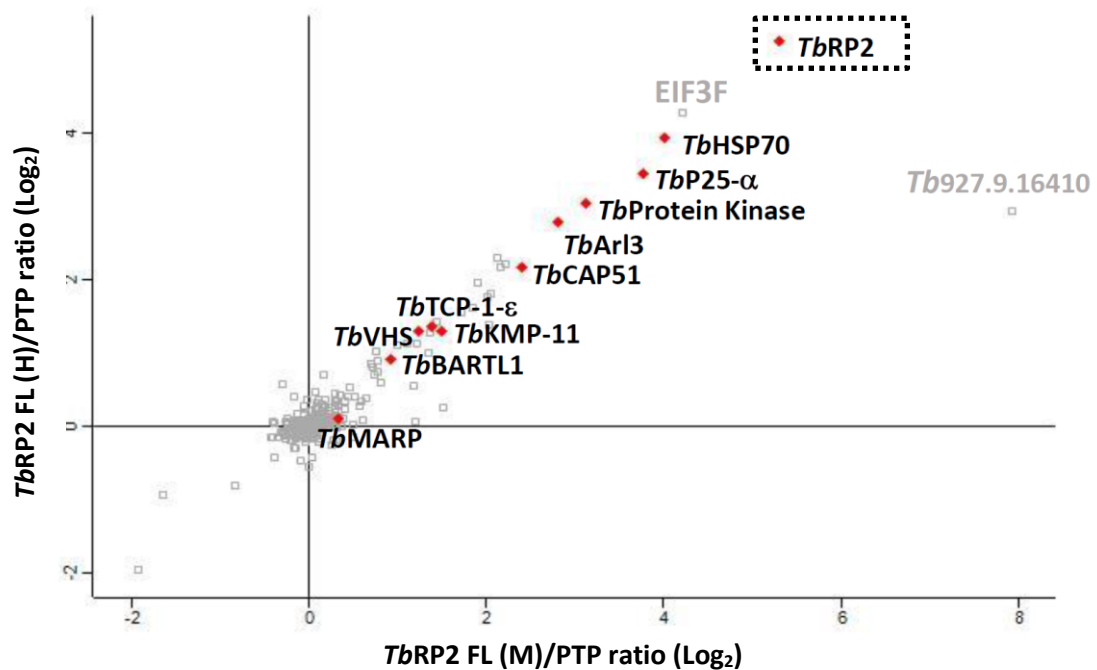
Tb927.10.14890	C-terminal motor kinesin, putative (TBKIFC1)	1.418	1.774	6.785	6.299	6.291	6.332	90.8	6;6	6;6	10.9
Tb927.10.12860	hypothetical protein, conserved (pre-RNA processing PIH1/Nop17, putative)	2.012	1.761	6.393	5.562	5.878	6.132	77.741	3	3	5.8
Tb927.11.1980 (previously known as : Tb11.46.0009)	zinc finger protein family member, putative (ZC3H41)	1.856	1.632	7.068	6.330	6.611	6.738	58.459	6	6	13.3
Tb927.5.3120	eukaryotic translation initiation factor 2 beta subunit, putative (EIF2B)	1.710	1.563	6.479	5.744	6.027	6.145	34.967	1	1	5.4
Tb927.9.12650 (previously known as : Tb09.211.3610)	ubiquitin-activating enzyme E1, putative (UBA2)	1.434	1.436	6.994	6.300	6.547	6.637	134.65	7;2	7;2	7.7
Tb927.10.4430;Tb10.v4.0033;Tb11.v5.0767	pumillo RNA binding protein PUF1 (PUF1)	2.031	1.399	6.491	5.556	6.310	5.843	63.851	3;3;2	3;3;2	7.2
Tb927.9.13920;Tb927.9.13880;Tb927.9.13820	kinetoplastid membrane protein KMP-11 (KMP-11)	1.379	1.362	7.827	7.155	7.376	7.464	11.076	7;7;7	7;7;7	60.9
Tb927.11.14250	T-complex protein 1, epsilon subunit, putative (TCP-1-epsilon)	1.243	1.309	7.021	6.330	6.602	6.639	59.381	7	7	13.9

Tb927.11.10770	VHS domain containing protein, putative	1.498	1.308	6.031	5.160	5.576	5.743	48.943	1	1	3.7
Tb927.4.2000	ruvB-like DNA helicase, putative, ATP-dependent DNA helicase, putative	1.364	1.288	6.396	5.718	5.930	6.048	52.61	3	3	8.4
Tb927.6.400	peptidase M20/M25/M40, putative	1.226	1.140	7.457	7.325	6.534	6.613	52.185	5	5	13.5
Tb927.11.13050; Tb927.11.13040; Tb927.11.13030; Tb927.11.13020	Calmodulin	1.116	1.139	6.808	6.244	6.288	6.437	16.838	4;4;4;4	4;4;4;4	38.9
Tb927.5.930; Tb11.v5.0613	NADH-dependent fumarate reductase (FRDg)	1.006	1.102	7.151	6.562	6.673	6.764	123.69	10;6	3;0	3.1
Tb927.7.6090	hypothetical protein, conserved	0.756	1.036	6.971	6.383	6.499	6.578	88.162	8	8	13
Tb927.10.1500	methionyl-tRNA synthetase, putative (MetRS)	1.350	1.000	6.977	6.382	6.519	6.576	86.91	7	7	11.6
Tb927.10.5810	BARTL1, putative	0.926	0.930	6.210	5.561	5.705	5.875	48.257	2	2	7.2

Table 3.4 Proteins identified in *TbRP2* BirA\* fusion protein expressing *T. brucei* by BioID-SILAC (WC-2). Yellow highlight indicated bait protein -*TbRP2*, and proteins highlighted in blue were selected for further analysis. M/L :*TbRP2* FL/PTP; H/L :*TbRP2* FL /PTP.

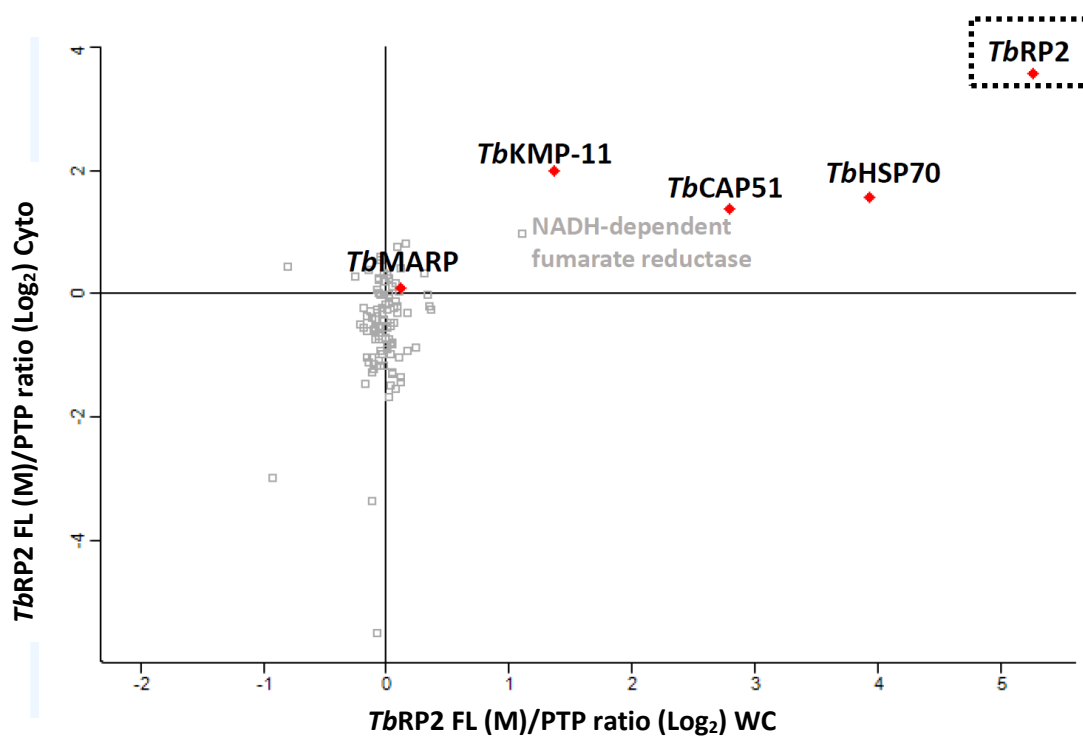


**Figure 3.21** Scatter plot showing proteins identified in the SILAC-enriched 'medium'-labelled *TbRP2* FL-BirA\* cell line (upper panel) or 'heavy'-labelled *TbRP2* FL-BirA\* cell line (lower panel). The x-axis shows the log2 normalised ratio of 'medium or heavy' labelled proteins versus 'light'-labelled proteins, and the y-axis represents the relative abundance of each protein.



**Fig 3.22** The LC-MS/MS scatter plot comparing proteins identified in the SILAC-enriched ‘medium’-labelled *TbRP2 FL::myc::BirA\** cell line ratio versus ‘light’-labelled PTPs parental cell line against proteins identified in ‘heavy’-labelled *TbRP2 FL::myc::BirA\** cell line ratio versus ‘light’-labelled PTPs parental cell line. The x-axis shows the ratio of ‘medium’-labelled proteins versus ‘light’-labelled proteins, identified according to the peptides detected by mass spectrometry. The y-axis shows the ratio of ‘heavy’-labelled proteins versus ‘light’-labelled proteins, identified according to the peptides detected by mass spectrometry. Scales is normally distributed and the ratios normalised because the base peak originally lay to the left of zero. Proteins which appear in the centre are more likely to be contaminant.





**Fig 3.23** The LC-MS/MS scatter plot comparing proteins identified in the SILAC-enriched 'medium'-labelled *TbRP2 FL::myc::BirA\** cell line ratio versus 'light'-labelled PTPs parental cell line against proteins identified in 'heavy'-labelled *TbRP2 FL::myc::BirA\** cell line ratio versus 'light'-labelled PTPs parental cell line. The x-axis shows the ratio of 'medium'-labelled proteins versus 'light'-labelled proteins, identified according to the peptides detected by mass spectrometry. The y-axis shows the ratio of 'heavy'-labelled proteins versus 'light'-labelled proteins, identified according to the peptides detected by mass spectrometry. Scales is normally distributed and the ratios normalised because the base peak originally lay to the left of zero. Proteins, which appear in the centre are more likely to be contaminant.

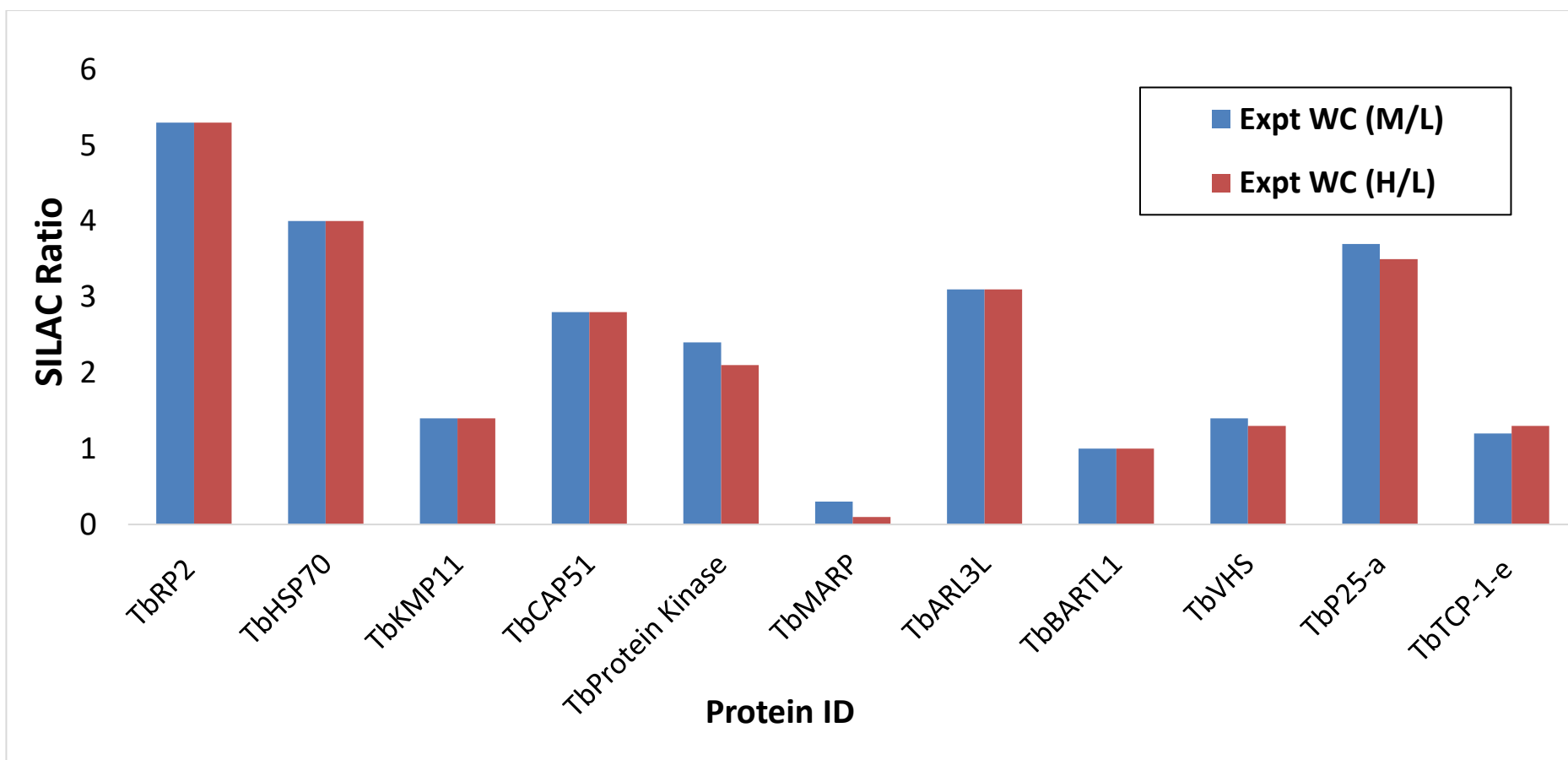


Fig 3.24 The SILAC ratio of each selected proteins in the BioID whole cell fraction is indicated by the colour of each bar

### 3.8 Summary

BioID combined with SILAC was used to identify *TbRP2*-interacting and near neighbour proteins. Expression of *TbRP2* FL and *TbRP2*<sup>Δ134-463</sup> had no effect on cell growth or morphology. However, expression of *TbRP2*<sup>Δ322-463::myc::BirA\*</sup> and *TbRP2*<sup>Δ171-463::myc::BirA\*</sup> resulted in dominant negative phenotypes and so analysis of the cell lines expressing these constructs was not undertaken. A number of proteins were identified in the successful BioID-SILAC experiments, including a large number of proteins which were potentially contaminants e.g. ribosomal proteins; it was here that the use of SILAC was highly informative in sorting 'noise' from potentially 'real' *TbRP2* proximal proteins. The identification of *TbRP2* as greatly enriched in all the experiments indicated that the BioID approach was functioning correctly. Compared to other large proteomic studies the number of proteins identified was small (Urbaniak et al., 2013). In part, this was due to the fact that to get a SILAC ratio, a peptide has to be detected at least twice otherwise there would not be a SILAC ratio value. Therefore, it is possible that some 'real' *TbRP2*-interacting proteins will have fallen within the 'noise'. Nevertheless, a number of proteins of significance were identified for further analysis.

## Chapter 4 Bioinformatic validation

### 4.1 Introduction

In previous sections, I explained how BioID was used in conjunction with SILAC to identify putative *TbRP2* proximal proteins. In this section, I describe how bioinformatic approaches were used to prioritise proteins for further analysis. Firstly, all proteins with a SILAC ratio  $\geq 1$  were analysed by GENEDB and TriTrypDB to determine what was known about the proteins, including proposed GO function(s) and occurrence in other proteomic datasets. These initial steps, allowed me to gain some basic background information on the proteins identified from BioID. As the aim of my PhD was to interrogate the function(s) of the *TbRP2* protein in PCF *T. brucei*, proteins that might be associated with *TbRP2* functions were to be prioritised. For example, the BioID experiments identified several chaperone proteins, and proteins of high molecular weight, which could possibly, represented basal body/TZ components. Secondly, the amino acid sequence of each protein was interrogated using BLASTP, to identify the conservation between the *T. brucei* proteins of interest and homologous proteins in humans and other microbial eukaryotes (both flagellated/ciliated and non-ciliated eukaryotes were included in this organismal set Table 4.1).

In the following sections, the domain structure for each protein will be presented, along with amino acid sequence alignments of *T. brucei* proteins and putative orthologues from *Homo sapiens* and diverse microbial eukaryotes. It should be noted that some proteins with 'good' SILAC ratio i.e.  $\geq 1$ , were not selected for further analysis (Appendix CD for protein ID).

<b><u>Organisms</u></b>	<b><u>Builds cilia/flagella</u></b>
<i>Homo sapiens</i>	Y
<i>Trypanosoma brucei</i>	Y
<i>Trypanosoma cruzi</i>	Y
<i>Leishmania major</i>	Y
<i>Plasmodium falciparum</i>	Y
<i>Chlamydomonas reinhardtii</i>	Y
<i>Giardia lamblia</i>	Y
<i>Paramecium tetraurelia</i>	Y
<i>Caenorhabditis elegans</i>	Y
<i>Saccharomyces cerevisiae</i>	N
<i>Dictyostelium discoideum</i>	N
<i>Capsaspora owczarzaki</i>	N

Table 4.1 A broad range of microbial eukaryotes, including flagellated and non-flagellated organisms used in the protein analysis.

<b>Proteins</b> <b>Organisms</b>	<b>HSP70</b>	<b>KMP11</b>	<b>MARP</b>	<b>PK</b>	<b>CAP51</b>	<b>ARL3L</b>	<b>TPPP/ p25α</b>	<b>TCP-1-ε</b>	<b>VHS</b>	<b>TbBBP590</b>	<b>BARTL1</b>
<i>Homo sapiens</i>	✓	X	X	✓	X	✓	✓	✓	✓(N)	X	✓
<i>Trypanosoma cruzi</i>	✓	✓	✓ (only partial)	✓	✓	✓	✓	✓	✓	✓	✓
<i>Leishmania major</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<i>Plasmodium falciparum</i>	✓	X	X	✓	X	✓	✓	✓	X	X	X
<i>Chlamydomonas reinhardtii</i>	✓	X	X	✓	X	✓	✓	✓	X	X	✓
<i>Giardia</i> <i>Lamblia</i> (intestinalis)	✓	X	X	✓	X	✓	X	✓	X	X	X
<i>Paramecium tetraurelia</i>	✓	X	X	✓	X	✓	✓	✓	X	X	✓
<i>Caenorhabditis elegans</i>	✓	X	X	✓	X	✓	X	✓	✓(N)	X	X
<i>Saccharomyces cerevisiae</i>	✓	X	X	✓	X	✓	X	✓	X	X	X
<i>Dictyostelium discoideum</i>	✓	X	X	✓	X	✓	X	✓	X	X	X
<i>Capsaspora owczarzaki</i>	✓	X	X	✓	X	✓	X	✓	X	X	X

**Table 4.2 NCBI BLAST search of different organisms to the T. brucei BioID pull down proteins.** “✓” more than one BLAST hits ( $e^{-10}$  as cut-off point); “✓” only one BLAST hit and “X” no BLAST hits or very low e values. “✓” did not pass reciprocal BLAST and “X” no BLAST hits or very low e values ( $> e^{-10}$ ) (N) = only N-terminal amino acid sequence conserved between query and databases

<b><u>Proteins</u></b>  <b><u>Proteomic</u></b> <b><u>Studies</u></b>	<b>HSP70</b>	<b>KMP11</b>	<b>MARP</b>	<b>PK</b>	<b>CAP51</b>	<b>ARL3L</b>	<b>TPPP/ p25α</b>	<b>TCP-1-ε</b>	<b>VHS</b>	<b>TbBBP590</b>	<b>BARTL1</b>
<b>Flagellum cytoskeleton<sup>1</sup></b> (Broadhead et al., 2006)	<b>Y</b>	<b>Y</b>	<b>N</b>	<b>N</b>	<b>N</b>	<b>N</b>	<b>N</b>	<b>N</b>	<b>N</b>	<b>N</b>	<b>N</b>
<b>Whole flagellum<sup>2</sup></b> (Subota et al., 2014)	<b>Y</b>	<b>N</b>	<b>Y</b>	<b>N</b>	<b>Y</b>	<b>Y</b>	<b>Y</b>	<b>Y</b>	<b>Y</b>	<b>N</b>	<b>Y</b>
<b>Transition zone<sup>3</sup></b> (Dean et al., 2016)	<b>Y</b>	<b>Y</b>	<b>Y</b>	<b>Y</b>	<b>Y</b>	<b>Y</b>	<b>N</b>	<b>Y</b>	<b>N</b>	<b>Y</b>	<b>N</b>

**Table 4.3 The *T. brucei* BioID pull down proteins compares with proteomic studies**

(1) Broadhead R, et al. (2006) Nature; 440(7081):224–227; (Flagellar proteome)

(2) Subota I, et al. (2014) Mol Cell Proteomics; 13(7):1769–1786; (Intact flagella of procyclic Trypanosoma brucei)

(3) Dean S, et al (2016) Proc Natl Acad Sci; 113(35): E5135-43; (Transition zone)

## 4.2 Heat shock protein 70

Proteins belonging to the heat shock protein 70 (HSP70) family have been highly conserved through evolution; being found in both prokaryotic and eukaryotic organisms (Daugaard et al., 2007). The HSP70 protein was first thought to be involved in a response to stress (Pelham, 1984), but later studies indicated that HSP70 might play distinct roles within different organisms (Daugaard et al., 2005). This concept was based on experiments performed in mice, where the deletion of different Hsp70 genes led to remarkably discrete phenotypes. For example, among the members of HSP70 family; Hsc70, Hsp70-1t and Hsp70-2 are not stress-inducible proteins. Hsc70 can be found in cytosol and nucleus (Dworniczak and Mirault, 1987), and due to the essential role of Hsc70 for cell survival, a Hsc70 knockout cannot be generated (Daugaard et al., 2005; Florin et al., 2004). RNAi knockdown of Hsc70 results in massive cell death not only in various cancer cell types (such as HeLa (cervix), MCF-7-S1 and MDA-MD-468 (breast), PC-3 (prostate), HuH-7 (liver), and LoVo-36 (colon)), but also in non-tumorigenic cells. However, only tumorigenic cell growth depended on Hsp70 and Hsp70-2 (Rohde et al., 2005). Both Hsp70-1t and Hsp70-2 are also expressed in most of the tissues, but reported high levels in brain and testis (for Hsp70-1t) respectively (Daugaard et al., 2007; Goate et al., 1987; Son et al., 2000). In male mice, knockout of Hsp70-2 can result in massive germ cell apoptosis (Dix et al., 1996). Furthermore, Hsp70-2 has also been shown to be involved in the formation of the cyclin B/cdc2 complex during meiotic cell division and HSP70-2 knockout mice fail to complete meiosis I (Dix et al., 1997; Zhu et al., 1997). Other fundamental functions of HSP70 for maintaining of cellular homeostasis, and coordinating a number of essential cellular processes via folding (newly synthesised) and refolding (misfolded) proteins, transport proteins between cellular compartments, control of regulatory proteins and preventing denatured or unstable proteins (Kabani and Martineau, 2008; Mayer and Bukau, 2005).

Within the HSP70 protein family, variations in motif/domain architecture give rise to a variety of protein isoforms. However, here I focus only on *Tb*HSP70 (*Tb*927.11.11330) and its homologous sequences in *T. brucei* and human. There are

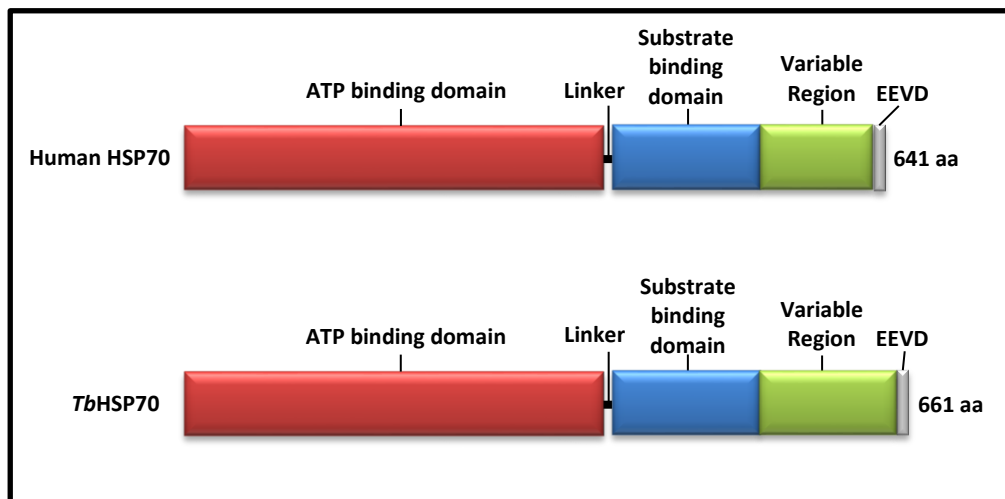


two major families of the eukaryotic HSP70 superfamily, (1) HSPA, which includes 13 of typical HSP70 proteins, and (2) HSPH that consist of four members of Hsp110/Grp170 proteins in human (Kampinga et al., 2009; Vos et al., 2008). A number of proteomic studies have shown that most of the *T. brucei* HSP70 proteins are expressed in both bloodstream and procyclic form parasites (Berriman et al., 2005; Broadhead et al., 2006; El-Sayed et al., 2005; Glass et al., 1986; Jones et al., 2006; MacFarlane et al., 1990; Subota et al., 2014; Vertommen et al., 2008). This might indicate that *TbHSP70* is likely to be involved in adapting parasites for different environment, and also suggests that *TbHSP70* might be important for parasite differentiation (Daugaard et al., 2007; Folgueira and Requena, 2007). A recent study by Drini and co-workers combined phylogenetic and comparative analyses of the TriTryp genomes, the draft genome of *Paratrypanosoma* (Flegontov et al., 2013), and the recently published genome sequences of 204 *L. donovani* field isolates (Imamura et al., 2016). This study provided novel understanding of how *Leishmania* HSP70 proteins may be involved in a response to specific environmental changes over evolution (Drini et al., 2016).

At the starting point of my analysis, I performed a BLAST search in which I used *TbHSP70* (*Tb927.11.11330*) as my query against the predicted protein database for *Homo sapiens* and all of the organisms listed in Table 4.2. NCBI BLAST search revealed that there are more than one HSP70 proteins, which above the cut-off  $e^{-10}$  value in every organism listed in Tble 4.2. In humans, the top protein hit is HSPA1A (also called inducible HSP70 or HSP72) (NP\_005337.2). This protein has been well studied and is highly stress-inducible (Chen et al., 2013; Liu et al., 2012; Shi et al., 1998; Zhang et al., 2014). I also performed a reciprocal BLAST using HSPA1A as the query sequence to check whether there are other HSP70 proteins in *T. brucei* that are more similar to the human HSPA1A sequence. This reciprocal BLASTP approach revealed *TbHSP70* (*Tb927.11.11330*) as the top hit; with an expectancy value of 0. Fig. 4.1-A displays the domain architectures of human and *TbHSP70* proteins. The human HSP70 is a 641 amino acid long protein with several domains. An ATPase domain - when ATP is hydrolysed to ADP it can lead to conformational changes in the second domain called substrate/peptide domain to trigger the release of substrate. There is a short linker

segment (~7 amino acids) between these two domains. At the C terminus of the HSP70 is a G/P-rich region, which contains an EEVD-motif that enables proteins to bind co-chaperones such as HSP40 and HSP70/HSP90 organising proteins (HOP) (Goloubinoff, 2017; Louw et al., 2010; Mayer et al., 2001; Mayer and Bukau, 2005; Odunuga et al., 2004). A similar domain structure can be found in *Tb*HSP70 and the sequence alignment shows high degree of conservation between the different domains highlighted in different colours (Fig. 4.2-B).

**(A)**



**(B)**

H.sapiens TbHSP70	<b>MAKAAAIGIDLGTTYSCVGVFQHGKVEIIANDQGNRTTPSYVAFTDTERLIGDAAKNQVA</b> MTYEGAIGIDLGTTYSCVGVWQNERVEIIANDQGNRTTPSYVAFTDSERLIGDAAKNQVA *: .*****: *: :*****:*****
H.sapiens TbHSP70	<b>INPQNTVFDKRLIGRKFGDPVVQSDMKHWPQVINDG-DKPKVQVSYKGETKAFYPPEEI</b> MNPTNTVFDKRLIGRKFSVQSDMKHWPQVVTGDDKPVIQVQFRGETKTFNPPEEI :* *****.* *****:*. *. * * :*: :*: * * *
H.sapiens TbHSP70	<b>SSMVLTKMKEIAEAYLGYPVTNAVITVPAYFNDSQRQATKDAGVIAGLNLVRIINEPTAA</b> SSMVLTKMKEVAESYLGKQVAKAVVTVPAYFNDSQRQATKDAGTIAGLEVLRIINEPTAA ***** * *: *: * *: :*: :*****. *: :*****
H.sapiens TbHSP70	<b>AIAYGLDRTGKG-ERNVLIFDLGGGTFDVSILTIDDGIFEVKATAGDTHLGGEDFDNRLV</b> AIAYGLDKADEGKERNVLIFDLGGGTFDVTLLTIDDGIFEVKATNGDTHLGGEDFDNRLV *****: :.* *****: :*. ***** *****
H.sapiens TbHSP70	<b>NHFVEEFKRKHK-KDISQNKRAVRLRTACERAKRTLSSSTQASLEIDSLFEGIDFYTSI</b> AHFTEEFKRKNKGDLSNLRALRLRTACERAKRTLSSAAQATIEIDALFENIDFQATI *. *****.* *: *. * *:*****: :*: :*: * * *: *
H.sapiens TbHSP70	<b>TRARFEELCSDLFRSTLEPVEKALRDAKLDKAQIHDLVLVGGSTRIPKVQKLLQDFNGR</b> TRARFEELCGDLFRGTLQPERVLQDAKMDKRAVHDVVLVGGSTRIPKVMQLVSDFFGGK *****.*****. *: *: :*: :*: * * :*:***** :*: * * *: *
H.sapiens TbHSP70	<b>DLNKSINPDEAVAYGAAVQAAILMGDKSENVQDLLLLDVAPLSLGLTAGGVMTALIKRN</b> ELNKSINPDEAVAYGAAVQAIFLTGGKSKQTEGLLLLDVAPLTGLIETAGGVMTALIKRN :***** ***** * *. *: :. *****. *: *****
H.sapiens TbHSP70	<b>STIPTKQTQIFTTYSNQPGLIQVYEGERAMTKDNLLGRFELSGIPPAPRGVPQIEVT</b> TTIPTKKSQIFSTYSNQPGLVHIQVFEGERMTKDCHELLGTGFDLSGIPPAPRGVPQIEVT :*****: :*: :***** * *:*****: * * *: * *:*****
H.sapiens TbHSP70	<b>FDIDANGILNVTATDKSTGKANKITITNDKGRLSKEEIERMVQEAKEYKADEVQRRVS</b> FDLDANGILSVSAEEKGTGKRNVITNDKGRLSKADIERMVSDAAKYEADKAQRRERID *:*****.* *: * * *:***** :*****.* *:*****.*****:
H.sapiens TbHSP70	<b>AKNALESYAFNMKSAVEDEGLKGKISEADKKKVLDKCQEVI SWLDANTLAEKDEFEHKRK</b> AKNGLENYAFSMKNTINDPNVAGKLDADKNAVTTAVEEALRWLNDNQEASLDEYNHRQK **.*.*.*.*.:*: :*: :* *: * * *:*****
H.sapiens TbHSP70	<b>ELEQVCNPIISGLYQGAGGPGPGG-----FGAQGPKGGSGSGPTIEEV</b> ELEGVCAPILSKMYQGMGGDAAGGMPGGMPGGMPGGMGGMGGAASSGPKVEEV *** * *: :* * * . * :* . . . :*
H.sapiens TbHSP70	D D *

**Figure 4.1 (A) Schematic of human HSP70 (NP\_005337.2) and *T. brucei* HSP70 (Tb927.11.11330) domain structure.** Different domains and co-chaperone-binding sites are indicated by different colours. Typical cytosolic human HSP70 proteins with 3 domains: The N-terminal ATPase binding domain, a peptide (substrate) binding domain and characteristic C-terminal EEVD motif (can bind TPR-domain-containing proteins). (B) Multiple sequence alignments between those two HSP70 proteins by ClustalOmega analysis. Two HSP70 proteins are very well conserved. “\*” indicates identical aa and “:” indicates conserved, **red highlights = ATPase binding domain**; **blue highlights = substrate binding domain**; **green highlights = variable region** and black highlights = linker structure between the domains.

When I performed a BLAST search using *TbHSP70* (*Tb927.11.11330*) as my query against *T. brucei* itself, 11 proteins were identified with an e value above the  $e^{-10}$  cut-off point (see table 3.5). The cartoons of all 12 HSP70 sequences and protein domains are presented in Fig 4.2 A-B; showing the proteins are well conserved in some regions (Appendix II Fig 9.5 sequence alignments). Most of the *T. brucei* HSP70 proteins are expressed in both bloodstream and procyclic form, and they can be further divided into four subdivisions based on their localisations, which are cytoplasmic HSP70s, mitochondrial HSP70s, endoplasmic reticulum HSP70s and other HSP70s. Based on their intracellular localisations, different HSP70s could serve different purposes (Jones et al., 2006; Louw et al., 2010).

### **Cytoplasmic HSP70s**

In *T. brucei* there are four typical HSP70 proteins, including the query sequence *TbHSP70* (*Tb927.11.11330*), which are predicted to have cytoplasmic localisations. The other proteins are *TbHSP70.4*, *TbHSP70.c*, and *TbHSP110*. A sequence alignment of these four HSP70 proteins indicated a high degree of amino acid sequence conservation. The alignment shows a key difference between cytoplasmic *TbHSP70* and the rest of the family is that the former, but not the latter, contains the C-terminal EEVD motif. *TbHSP70.c* has been identified as a cytosolic HSP70 protein (Burger et al., 2014). *In silico* analysis indicates that *TbHSP70.c* consists of atypical acidic residues in the substrate binding domain, and a RIEAINANTE sequence at the C-terminus of the protein, rather than the EEVD motif (Burger et al., 2014). *Tbj2* protein (*Tb927.2.5160*), is one of the six *T. brucei* Type I Hsp40s, and evidence indicates that *Tbj2* is a potential co-chaperone partner for *TbHSP70.c*, which can stimulate of *TbHsp70.c* ATPase activity and prevent thermal aggregation of malate dehydrogenase (MDH) and rhodanese (Ludewig et al., 2015). *TbHSP70.4* is very closely related to so-called typical *TbHSP70*. *TbHSP70.4* and *TbHSP70* share 64% of identity at an amino acid sequence level, which suggests they might be functionally related. After HSP70 and HSP90, HSP110 is the third most abundant heat shock proteins in most all mammalian organisms (Easton et al., 2000; Lee-Yoon et al., 1995). Members of the HSP110 family

can also be found in plants and fungi as well (Easton et al., 2000). The exact role(s) of HSP110 is unclear, but some studies have suggested that HSP110 is involved in preventing protein aggregation and temperature resistance (Raviol et al., 2006), and further study has shown that HSP110 can act as a nucleotide exchange factors (NEFs) of HSP70s in the cytosol of many organisms (Andréasson et al., 2008). From an amino acid alignment, the *T. brucei* HSP110 shares good conservation at the N-terminal ATPase binding domain region with HSP70 but lacks similarity at the C-terminal domain region. So far, as with *TbHSP70.4*, little is known about HSP110 in *T. brucei*, however, HSP110 in the closely related kinetoplastid; *Leishmania infantum*, has been identified as a amastigotes specific protein (McNicoll et al., 2006).

### **Mitochondrial HSP70s**

*T. brucei* possess three mitochondrial Hsp70 proteins - mt*TbHSP70* protein (A-C), which share good conservation to typical *TbHSP70* (Louw et al., 2010). All three mt*TbHSP70* proteins possess ATPase binding domain that contains about 409 amino acids, and a partially conserved linker and substrate binding domain region at the C terminus of the sequence that comprises ~ 242 amino acids, but lacks the EEVD motif. Based on proteomic studies, mt*TbHsp70A* and mt*TbHsp70C* are expressed in both BSF and PCF, while mt*TbHsp70B* can only be detected in the membrane fraction of BSFs (Jones et al., 2006; Vertommen et al., 2008). Interestingly, three copies of mtHsp70 in the *T. brucei* genome are located on chromosome 6 and are 100% identical at the amino acid level. The presence of two identical mtHsp70s (mt*TbHsp70A* and mt*TbHsp70B*) in the BSF stage, could be that in mammalian host; the only food for the parasites in the mammalian host is glucose, and the level of ATP is much lower than in PCFs. Therefore, they have to keep the amount of mitochondrial machinery and function in the mammalian host low (Williams et al., 2008). The mtHSP70 proteins have been shown play multiple roles, such as cooperating with the mitochondrial nucleotide exchange factor mGrpE (Mge1p) in both importing and folding of mitochondrial proteins (Azem et al., 1997). Later studies demonstrated additional roles for mtHsp70 protein in *T. brucei*, including involvement in importing tRNA into

the mitochondrion (Tschopp et al., 2011) and its participation in kDNA replication and maintenance (Týč et al., 2015).

### Endoplasmic reticulum HSP70s

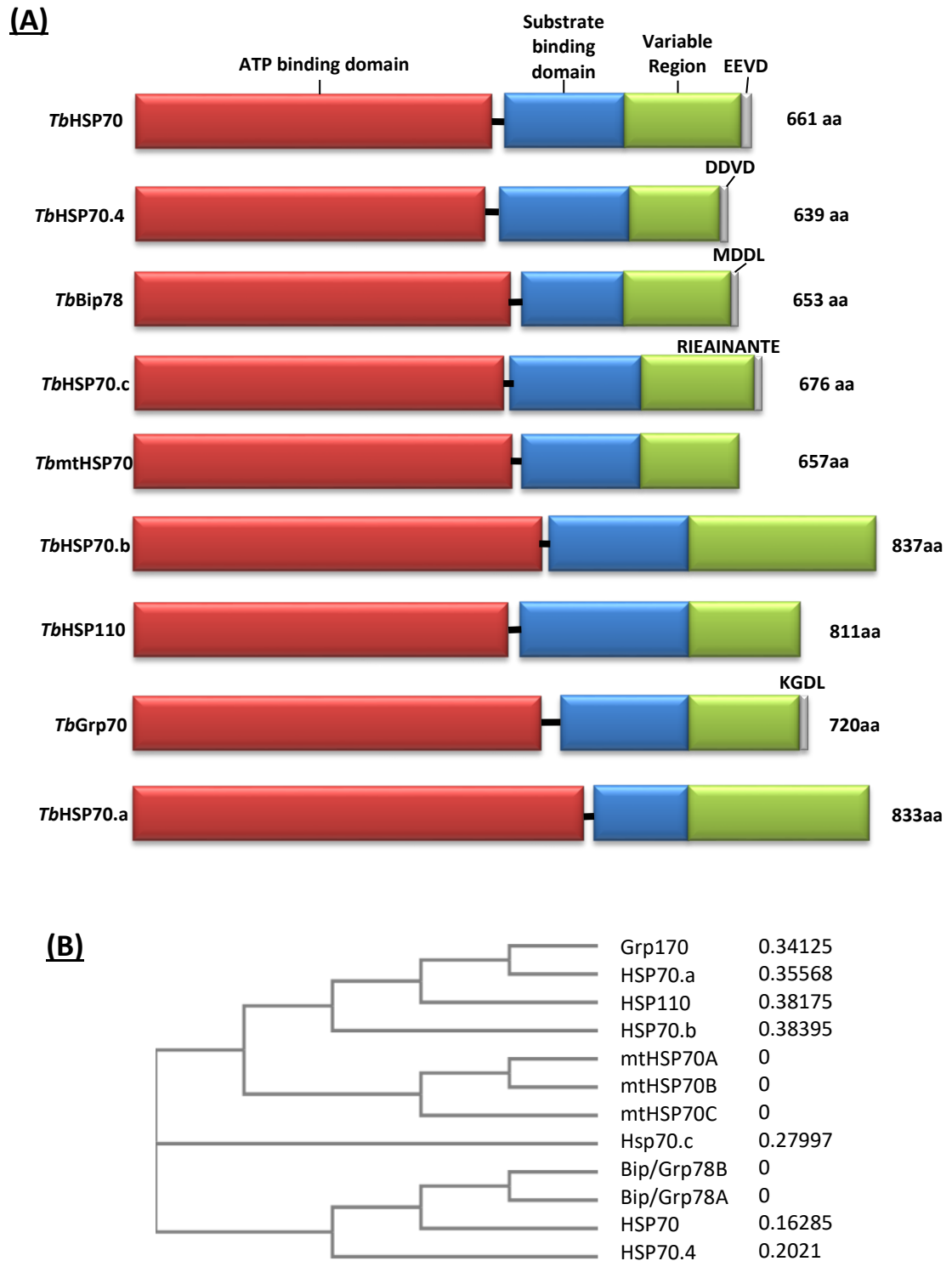
ER-HSP70s can be divided in two groups: (1) BiP/Grp78 (**B**inding **i**mmunoglobulin **p**rotein/**G**lucose-**r**egulated **p**rotein), which are both expressed in BSF and PCF (2) Grp170 that only can be found in BSF membrane. The possible function(s) of those proteins play important roles in protein folding, post-translational modification, maintain cellular homeostasis under stress conditions (Dudek et al., 2009). The two genes *TbBiP/Grp78A* and *TbBiP/Grp78B* are identical but are transcribed separately on chromosomal 11 (Louw et al., 2010). The level of *TbBiP/Grp78* proteins are much higher in BSF than PCF, which could be due to the level of VSG abundance that needs to be maintained (Bangs et al., 1993). The *TbGrp170* protein belongs to ER-resident member of HSP110 family and homologous genes can be found in eukaryotes, but not prokaryotes. However, the function(s) of this protein still need to be elucidated.

### Other HSP70s

Bioinformatics analysis has identified two novel HSP70s in the *T. brucei* genome - *TbHSP70.a* and *TbHSP70.b*. The exact functions of both proteins are still under investigation. However, the gene that encoded *TbHSP70.b* is flanked by genes encoding 40S ribosomal proteins, which may suggest that *TbHSP70.b* is involved in ribosomal assembly. Based on the phylogenetic tree generated by Clustal Omega, both HSP70.a and HSP70.b are more closely related to the Hsp110/Grp170 family, and HSP70.b in particular forms a well-supported distinct group within this family Fig 3.28 C. Furthermore, the *TbHSP70.b* is more similar to cyanobacterial DnaK than the rest of the Hsp110/Grp170 families, which might suggest that *TbHSP70.b* protein has a bacterial origin (Folgueira and Requena, 2007).

Protein ID	Accession Number	e Value	Life stages	Localisations
HSP70	<i>Tb</i> 927.11.11330 ( <i>Tb</i> 11.01.3110)	0	BSF&PSF	Cytoplasm
HSP70.4	<i>Tb</i> 927.7.710	0	BSF&PSF	Cytoplasm
BiP/Grp78B	<i>Tb</i> 927.11.7460 ( <i>Tb</i> 11.02.5450)	0	BSF&PSF	ER
BiP/Grp78A	<i>Tb</i> 927.11.7510 ( <i>Tb</i> 11.02.5500)	0	BSF&PSF	ER
HSP70.c	<i>Tb</i> 927.11.11290 ( <i>Tb</i> 11.01.3080)	0	BSF&PSF	Cytoplasm
Mitochondrial HSP70A	<i>Tb</i> 927.6.3740	0	BSF&PSF	Mitochondria
Mitochondrial HSP70B	<i>Tb</i> 927.6.3750	0	BSF&PSF	Mitochondria
Mitochondrial HSP70C	<i>Tb</i> 927.6.3800	0	BSF&PSF	Mitochondria
HSP70.b	<i>Tb</i> 927.7.1030	4E <sup>-45</sup>	Unknown	Nucleoplasm ?
HSP110	<i>Tb</i> 927.10.12710 ( <i>Tb</i> 10.389.0880)	1E <sup>-40</sup>	BSF&PSF	Cytoplasm
Grp170	<i>Tb</i> 927.9.9860 ( <i>Tb</i> 09.211.1390)	3E <sup>-30</sup>	BSF	ER
HSP70.a	<i>Tb</i> 927.9.4500 ( <i>Tb</i> 09.160.3090)	2E <sup>-26</sup>	BSF	Unknown

**Table 4.4 The HSP70 family of *Trypanosoma brucei***



**Figure 4.2 (A)** Schematic cartoon of *T. brucei* HSP70s domain structure. Different domains and co-chaperone-binding sites are indicated by different colours. Typical cytosolic human HSP70 proteins with 3 domains: The N-terminal ATPase binding domain, a peptide (substrate) binding domain and characteristic C-terminal EEVD motif (can bind TPR-domain-containing proteins). “\*” indicates identical aa and “:” indicates conserved aa. **(B)** Phylogenetic analysis of the *TbHsp70* proteins.



### 4.3 ADP-ribosylation factor-like protein 3-Like (Arl3L) protein

Arl3 protein belongs to a member of the ADP-ribosylation factor family of GTP-binding proteins, which can be found in all ciliated organisms (Zhang et al., 2013). The protein cycles between an inactive Arl3-GDP-bound and an active Arl3-GTP-bound form. This interchange is tightly regulated by guanine nucleotide exchange factors (GEF) and GTPase-activating proteins (GAP) (Cavenagh et al., 1994). More recently, another small protein called ARL13b was identified from mice. Here it was proposed that ARL13b act as an GEF that can stimulate ARL3-GDP to exchange to Arl3-GTP (Gotthardt et al., 2015). In human, ARL3 localises to many cellular compartments, such as photoreceptor synaptic terminal, cell body, inner segment, and especially expressed strongly in connecting cilium of photoreceptor cells. Arl3 has also been shown to specifically interact with RP2 (i.e. RP2 is the GAP for ARL3) at the connecting cilium (Grayson et al., 2002). Additionally, in many other studies ARL3 also co-localised with UNC119 (Ishiba et al., 2007) and PDE $\delta$  (Zhang et al., 2007). Although, to date there is no evidence that mutations in ARL3 lead to any human ciliopathies. , in *Arl3*<sup>-/-</sup> mice, typical ciliopathic phenotypes can be observed e.g. impaired photoreceptor development, but also cysts in kidney, liver and pancreas (and animals die within 3 weeks of birth) (Schrack et al., 2006). There are three putative orthologues in trypanosomatids - ARL-3A (Sahin et al., 2004), ARL-3B and ARL3-like (identified in this work). Previous studies have shown: RNAi knockdown of *Tb*ARL-3A causes significant shortening of flagella in *T. brucei* (Sahin et al., 2004), and that overexpression of constitutively 'active' LdARL-3A (blocked under the GTP-bound form) results in very short flagella in *Leishmania donovani* (Cuvillier et al., 2000).

Like HSP70 protein, multiple ARL3 (like) genes can be found in all the organisms listed in Table 4.2, i.e. irrespective of whether they can produce cilia/flagella. At the start of my bioinformatics analysis, I performed a straightforward NCBI BLAST research (Altschul et al., 1990), using *Tb*ARL3-like as my query sequence against human genome databases. In this analysis human ADP-ribosylation factor-like protein 3

(NP\_004302.1) was the top hit. The sequence alignment reveals that *Tb*ARL3L shares good homology to human Arl3, especially within the ARF domain (Fig 4.3-A). I also checked by reciprocal BLAST if any other proteins, rather than *Tb*ARL3L, were more closely related to human ARL3. The results showed that *Tb*ARL3-A (*Tb*927.6.3450) was the top hit ( $4e^{-72}$ ), followed by *Tb*ARL3-B (*Tb*927.10.8580;  $1e^{-61}$ ), and *Tb*ARL2 (*Tb*927.10.4250;  $9e^{-55}$ ) and that *Tb*ARL3L (*Tb*927.6.3650) only comes out as the 4<sup>th</sup> top hit ( $2e^{-51}$ ). This result indicates that *Tb*ARL3L fails the reciprocal BLAST test, and that it is not the orthologue of human ARL3. Additionally, it might not be surprising that ARL2 also came back with a good e value, as ARL2 and ARL3 shares 53% identity and have a high degree of structural conservation in humans; with evidence suggesting that at a biochemical level Arl2 and Arl3 are remarkably closely related (Zhou et al., 2006). Actually, when I use human ARL3 as my query to check how many Arl3 genes in human, the result revealed that there are a number of Arl-like proteins with good percentage identity and e-value, but they are not Arl3. Based on the NCBI BLASTP results, there are proteins annotated as Arl3 or its isoforms. In *T. brucei*, ARL2 protein has shown involved in cytokinesis and depletion or overexpressing *Tb*ARL2 cause a defect in cleavage furrow, which could leads to cell cycle arrest and multinucleated cells (Price et al., 2010). *Tb*ARL3L is a 183 amino acid protein, with a small GTP-binding domain.

Next, I analysed both proteins using Interpro Scan or Prosite to look more closely at their domain architecture. Human ARL3 is a 182 amino acid protein with a ARF domain located between amino acids 11 to 177, and *Tb*ARL3L appears to have very similar domain structure as well (Fig 4.3-B). The amino acid alignments and domain architecture between three ARL3 proteins shows good conservation Clustal Omega and InterproScan analysis (Fig 4.4 A-B) (Goujon et al., 2010; Zdobnov and Apweiler, 2001)

(A)

Human-ARL3	MGLLSILRK--LSAPDQEVRIILLGLDNAGKTTLLKQLASEDIS-----HITPTQ
TbARL3L	----MLKGIRSRAKRDNEPRVLIVGLDNAGKTTVLNALGEDEVFVEGKVSHAAPEGPTQ *: : *:* *:***:*****:*: *.:::
Human-ARL3	GFNIKSVSQSQGFKLNVDIGGQRKIRPYWKNYFENTDILIYVIDSADKRKRFEETGGQELAE
TbARL3L	GFNIKTTLTRGNKRAKLCDLGGRALRDYWQDYYSNTDCIMYVVDSSDHRRLEESHAAFVD *****: . :: *:* ** *::*:** :*:*.**:****: :
Human-ARL3	LLEEKLSVCVPVLIFANKQDLLTAAPASEIAEGLNLHTIRDVRWVIQSCSALTGEGVQDG
TbARL3L	VLK--GIEGAPVLVFANKQDLATAKDAQAIAECLHLHDFDRDKWHIQGCSAKTGAGLEEG *: :. .***:***** ** *. *** *.** :*** *.**.*** ** *:::*
Human-ARL3	MNWVCKNVNAKKK
TbARL3L	VAWILSTCAP-- : *: ..

(B)



**Figure 4.3 (A)** Multiple sequence alignments between human ARL3 protein (NP\_004302.1) and *T. brucei* ARL3-like (*Tb927.6.3650*) by ClustalOmega analysis. **(B)** The ARF domains are in yellow highlights (B) Schematic cartoon of human ARL3 protein (NP\_004302.1) and *T. brucei* ARL3-like (*Tb927.6.3650*) domain structure. Two ARL3 proteins are well conserved (especially within the ARF domain). “\*” indicates identical aa and “:” indicates conserved, red highlights = ARF domain.

```

TbARL3L      ----MLKGIRSRAKRDNPRVLIVGLDNGAKTTVLNALGEDEVPEVGKVSAAPEGPTQ
TbARL3A      MGLLTLLRKL--RS_SDASPRILILGLDNAGKTSILRNLSGEDPT-----TTQATQ
TbARL3B      MGLLEFLIKI--RPFSRRTRRIIMLGLDNAGKTRLLRRICEEEVS-----DTFPTQ
              :*  :  *          *:::*****  :.  :  ::          **

TbARL3L      GFNIKTLTRGNKRAKLCDLGGQRALRDYWQDYYSNTDCIMYVVDSSDHRRL EESHAAAFVD
TbARL3A      GFNIKTVDCEGFKLNVWDIGGQKAIRAYWPNYFDEVDCLVYVVD SADKRR LDETA AELET
TbARL3B      GFNIQNI TADELK FVVVDVGGQKSLRSYWRHYFDHTDALVFVIDSADMERIEEARTELHY
              ****::      :  :  *.***::*  **  .*. .*.:::*.***.  .*:*:  :  :

TbARL3L      VL--KGIEGAPVLV FANKQDLATAKDAQAI AECLHLHDFRDRKWHI QGCSAKT GAGLEEG
TbARL3A      LLQEEKLREVPFLVLANKCDLATALSPEDISTALNLQNLRDRTWSIQKCSAKTGEGLQEG
TbARL3B      ILEEEKLVGVPLLLFANKQDIPEAASQEEVMSSLNLRDTINRPWHIELCSAETGEGLSSG
              :*  :  :  .*.:::***  *:  *  .  :  :  .*:*:  :  *  *  *:  ***. **  *. *

TbARL3L      VAWILSTCAP-----
TbARL3A      FMWAIKSIKK-----
TbARL3B      LSWVVDTLKKRRPSLRPDGQV
              .  *  :::

```

**TbARL3L** Small\_GTPase\_ARF 183aa

**TbARL3A** Small\_GTPase\_ARF 178aa

**TbARL3B** Small\_GTPase\_ARF 190aa

144

## 4.4 Kinetoplastid membrane protein-11 (KMP-11)

KMP-11 was originally identified as an abundant membrane protein in *Leishmania donovani*, which can promote T-lymphocyte proliferation (Tolson et al., 1994). KMP-11 is expressed in several other trypanosomatids including *Trypanosoma cruzi*, *Crithidia fasciculata*, *Leptomonas collosma* and *Phytomonas spp* (Stebeck et al., 1995). Data has shown that KMP-11 was more highly expressed in procyclic, epimastigote and metacyclic life cycle stages than in bloodstream forms (Stebeck et al., 1995). In the list of selected organisms (Table 4.2), KMP-11 was only expressed in kinetoplastid-specific species (with e value above  $e^{-10}$ ), and no orthologues of KMP-11 can be found in humans or any other organisms. I then took all the BLAST hits into InterproScan/Prosite to interrogate domain architecture. *T. brucei* is a 92 aa protein that belongs to the KMP-11 super family, with 5 identical KMP-11 genes found in *T. brucei*. By InterproScan analysis, all *T. brucei*, *T. cruzi* (TcCLB.510755.89) and *L. major* (LMJLV39\_350028900) KMP11 proteins are 92 aa long and the KMP11 superfamily designation the whole protein. When I used CLUSTAL OMEGA to align KMP-11 protein sequences, it revealed that the proteins are nearly identical and share > 90 % homology (Fig 4.5).

```

LmKMP11      MATTYEEFSAKLDRLDEEFNRKMQEONAKFFADKPDESTLSPMKHEHYEKFERMIKEHTE
TbKMP11      MATTYEEFAAKLDRLDAEFAKKMEEQNKRFFADKPDEATLSPMKHEHYEKFEKMIQEHTD
TcKMP11      MATTLEEFSAKLDRLDAEFAKKMEEQNKKFFADKPDESTLSPMKHEHYEKFEKMIQEHTD
          ****  ***:*****  **  :*:***  :*****:*****:***:***:

LmKMP11      KFNKKMHEHSEHFKQKFAELLEQQKAAQYPSK
TbKMP11      KFNKKMREHSEHFKAKFAELLEQQKNAQFPGK
TcKMP11      KFNKKMHEHSEHFKAKFAELLEQQKNAQFPGK
          *****:*****  *****  **:*.

```

**Figure 4.5 Multiple amino acid sequence alignments** between three KMP-11 Kinetoplastids orthologues; *TbKMP11* (Tb927.9.13920), *TcKMP11* (TcCLB.510755.89) and *LmKMP11* (LMJLV39\_350028900) with Clustal Omega using default parameters. “\*” indicates identical aa and “:” indicates conserved.

## 4.5 Microtubule-associated repetitive protein (MARP)

The microtubule-associated repetitive protein (MARP) encoded by *Tb927.10.10360* was identified as a potential interacting/nearby protein of *TbRP2* by the BioID-SILAC approach. This MARP was first identified in *T. brucei* in 1988 by Schneider and co-workers, who reported the protein consisted of more than 50 nearly identical tandem repeats with a periodicity of 38 amino acids, and localised along subpellicular microtubules. (Schneider et al., 1988). From Table 4.5, the results were only based on simple BLAST analysis. Apart from *T. brucei*, only closely related species (*T. cruzi* and *L. major*), return with a good e value (see the table below). When I performed InterproScan/Prosit for protein domain analysis, no putative conserved domains were detected. As a final step of the MARP analysis, I used CLUSTAL OMEGA to align the top hit proteins from *T. cruzi* and *L. major* with the *T. brucei* MARP. This showed that there is not much conservation between the three MARP sequences at the N-terminus, but some similarity at the C-terminus (Appendix II Fig 9.6-A). However, when comparing the two *Tb*MARPs sequences, they share a high degree of identity towards the C-terminus, and are almost identical within the last 308 aa (Appendix II Fig 9.6-B).

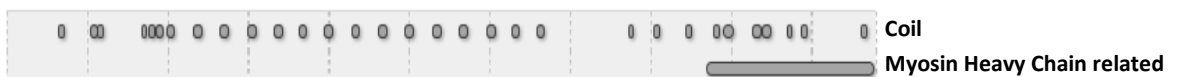
Protein ID ( <i>T. brucei</i> )	E-values	Protein ID ( <i>T. cruzi</i> )	E-values	Protein ID ( <i>L. major</i> )	E-values
Tb927.10.10360	0	TcCLB.511633.79 (Partial)	0	LmjF.05.0380	0
<i>Tb927.10.10280</i>	0	TcCLB.507447.19 (fragment)	0	LMJSD75_050008 700.1	1E <sup>-18</sup>
				LMJLV39_000005 900.1	3E <sup>-15</sup>

**Table 4.5** The orthologues of MARP (*Tb927.10.10360*) in *T. brucei*, *T. cruzi* and *L. major*.

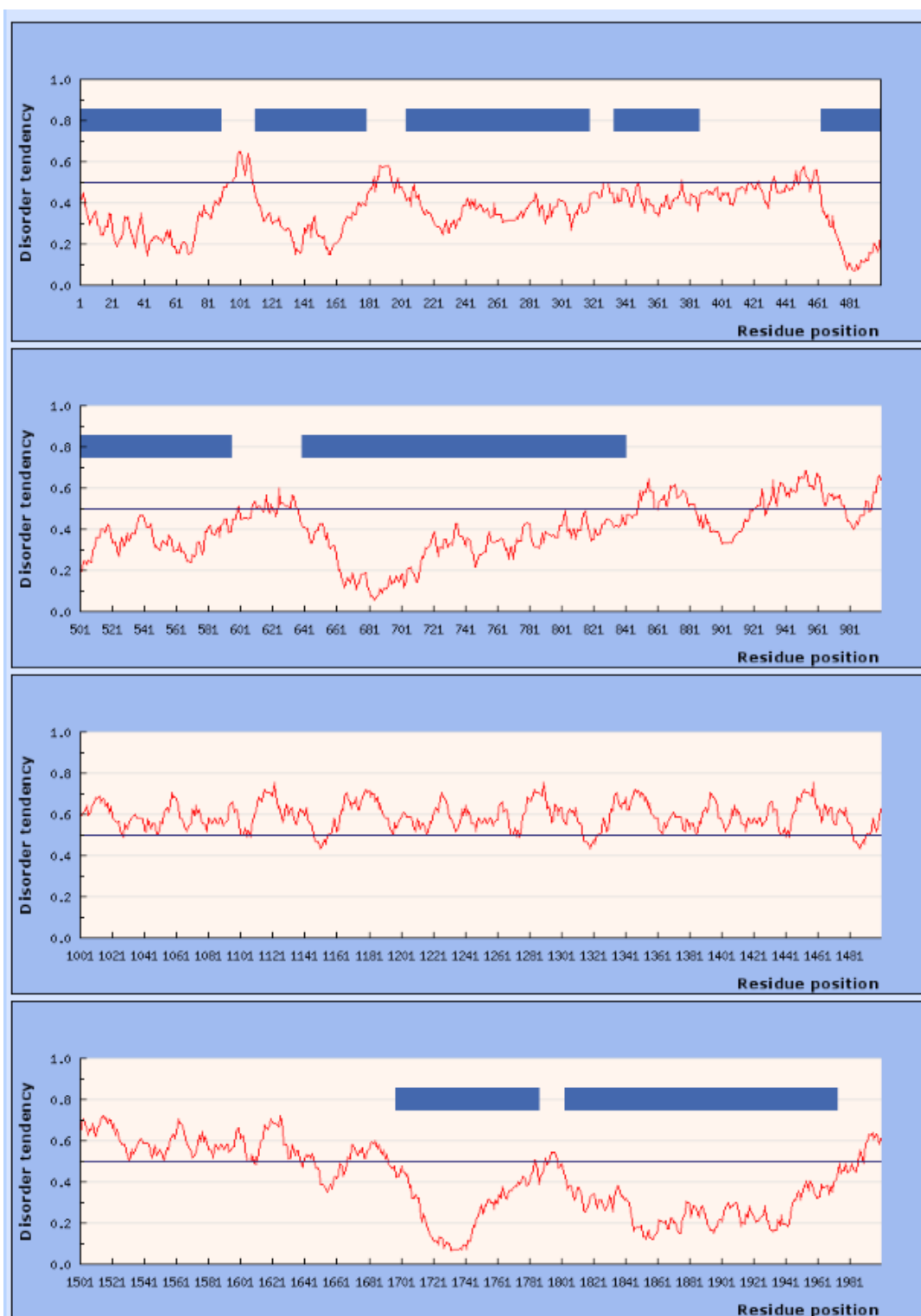
## 4.6 *Tb*BBP590

*Tb*BBP590 is another protein with high molecular weight (590.2kDa), and was identified very recently as a *T. brucei* basal body protein (McAllaster et al., 2015) (Dang et al., 2017). Based on the NCBI BLAST analysis, *Tb*BBP590 is a trypanosome-specific protein. Apart from *T. cruzi* (TCDM\_02199-t26) and *L. major* (LmjF.36.5250), sequences derived from all other organisms failed to return hits with a good expectancy value. The next step was to interrogate the domain structure of *Tb*BBP590. InterproScan and Prosite identified multiple coiled-coil regions (31 are predicted) and a myosin heavy chain-related domain located at the C-terminus. Additionally, IUPred (Dosztányi et al., 2005) was used here to predict any intrinsically unstructured/disordered and/or ordered/globular structure within the *Tb*BBP590 protein sequence. In this IUPred analysis scores above 0.5 indicate regions of disorder/unstructure and scores below 0.5 indicate that part of sequence has ordered structure (e.g. putative globular domain (1-89; 110-179; 204-318; 334-387; 463-595; 639-841; 1698-1787; 1803-1973; 3016-3105; 3121-3291; 3367-3633; 3662-3839; 3871-4021; 4038-4471; 4498-4648; 4856-5054; 5096-5228; 5260-5291)) (Fig 4.7-B). As a final step in the *Tb*BBP590 analysis, I used CLUSTAL OMEGA to align the *Tb*BBP590 sequences with *T. cruzi* (TCDM\_02199-t26) and *L. major* (LmjF.36.5250), which shows in Appendix II Fig 9.7. Compared to the *Tb*BBP590 protein sequence, both *T. cruzi* and *L. major* orthologues are much shorter. *T. cruzi* shares a good deal of conservation with *Tb*BBP590 towards the C terminus, but *L. major* is the opposite, showing good conservation with *Tb*BBP590 at the N-terminus (Appendix II Fig 9.7-A,B).

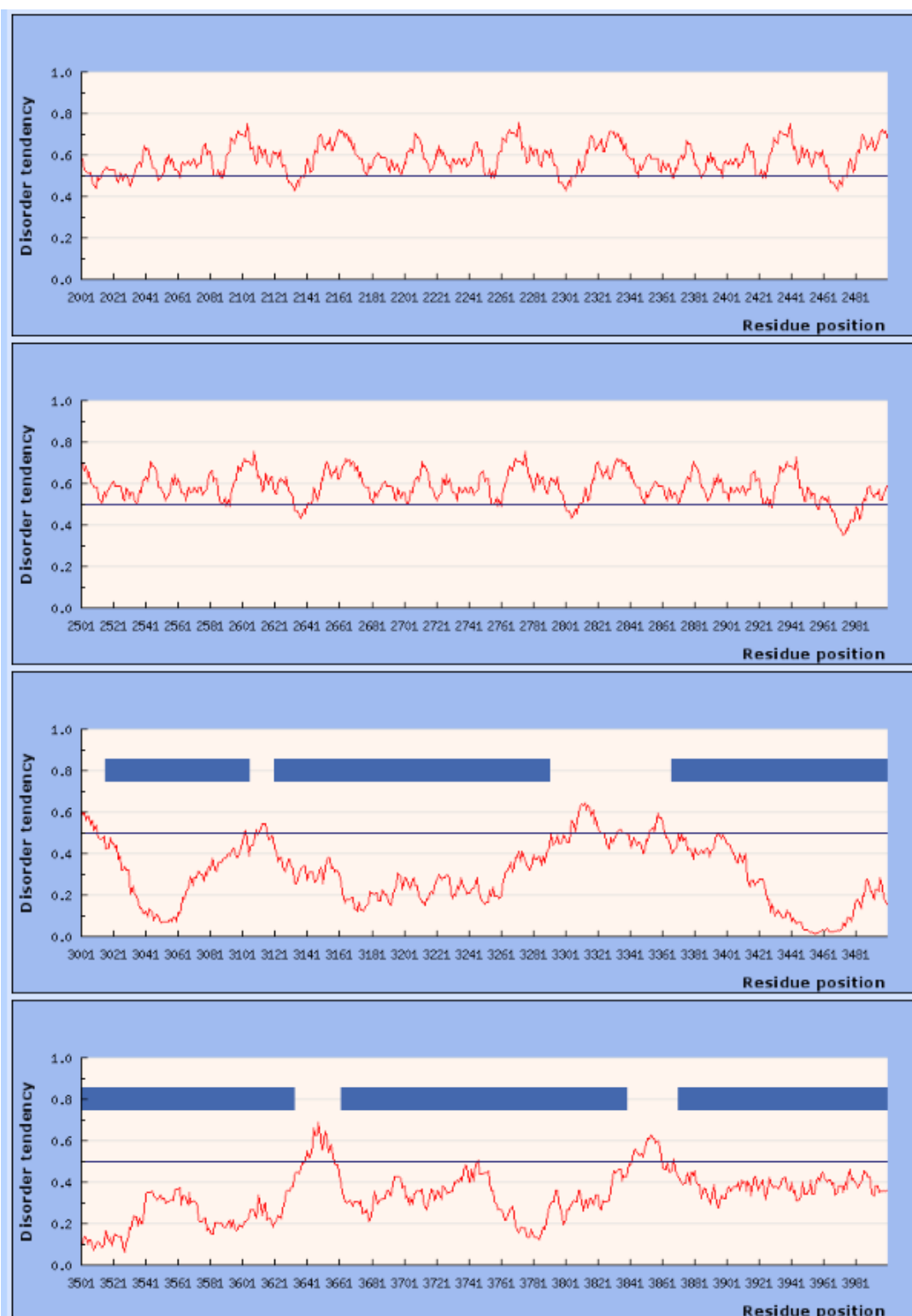
**(A)**

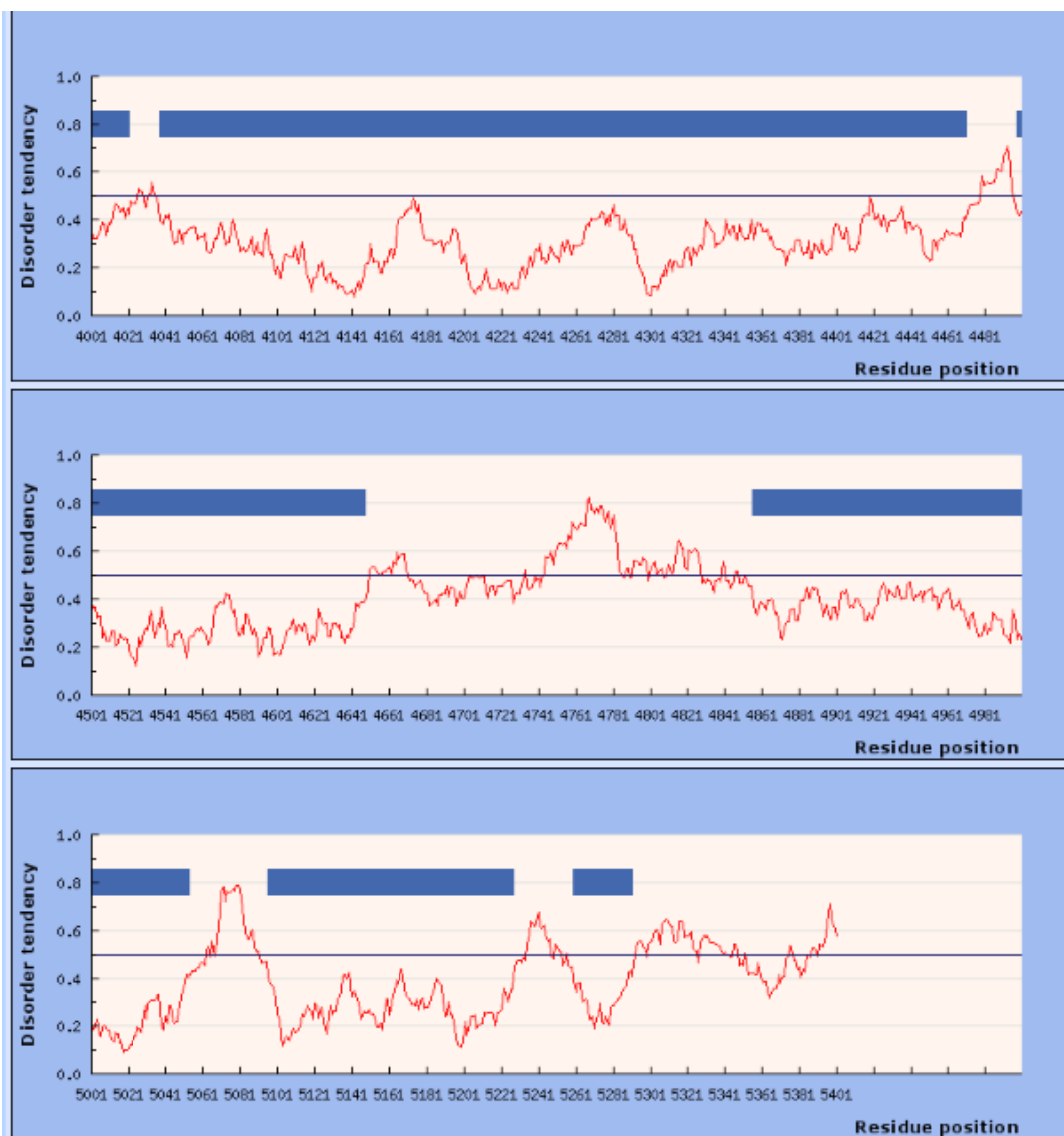


**(B)**







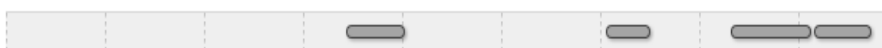


**Figure 4.7 (A)** Interpro Scan analysis of *TbBBP590*, a myosin heavy chain-related domain were detected at the C-terminus, and results also show possible coiled coil regions within the sequence. **(B)** Disorder prediction of structured regions performed by IUPRED

## 4.7 Cytoskeletal-Associated Protein (CAP) 51

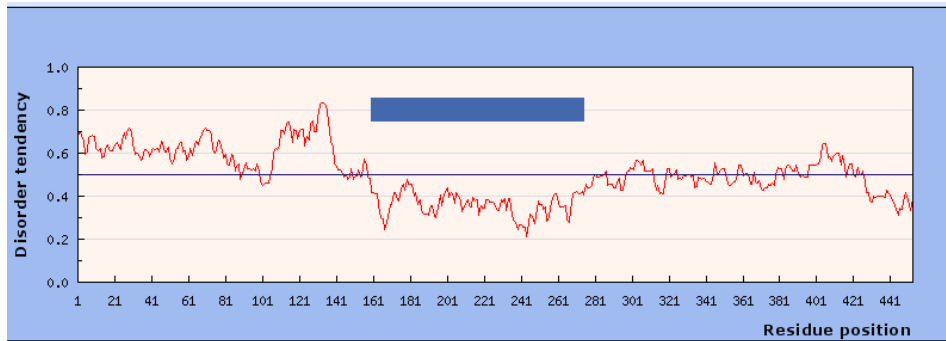
*TbCAP51* (Tb927.7.2640) along with another CAP protein *TbCAP51V* (Tb927.7.2650) were identified as a pair of paralogous proteins showing subpellicular microtubule localisation. However, *TbCAP51* is only expressed in the procyclic form, and *TbCAP51V* is detected only in the bloodstream form. This previous work also showed that both CAP51 and CAP51V play important roles in the organisation of the cytoskeleton and successful cytokinesis in their respective life cycle stages (Portman and Gull, 2014). There are no domain architectures in the *TbCAP51* protein that can be detected by either InterproScan or Prosite. However, as with *TbBB590*, the results did predict regions of coiled-coil within the sequence (Fig 4.8-A), and so I used IUPred and another tool called RONN (Yang et al., 2005) to predict possible regions of intrinsic protein disorder in *TbCAP51*. The IUPred analysis suggested that an intrinsically unstructured/disordered domain is likely at the N-terminus of the *TbCAP51* protein, and a putative globular domain located between 160-275 aa (Fig 4.8-B). Analysis using RONN also revealed a possible disordered region at the C-terminus (Fig 4.8-C). As a final step in the *TbCAP51* analysis, I used CLUSTAL OMEGA to align the sequences of *TbCAP51* with *TbCAP51V*, *T. cruzi* CAP (TcCLB.506859.180) and *L. major* (LMJLV39\_220012800) (Fig 4.8-D to E respectively). Compared to the *TbCAP51* amino acid sequence, *TbCAP51V* shares good conservation to *TbCAP51*, especially in the middle of the sequence where it is nearly identical. Both *T. cruzi* and *L. major* orthologues also show good conservation with *TbCAP51* at the C-terminus.

**(A)**

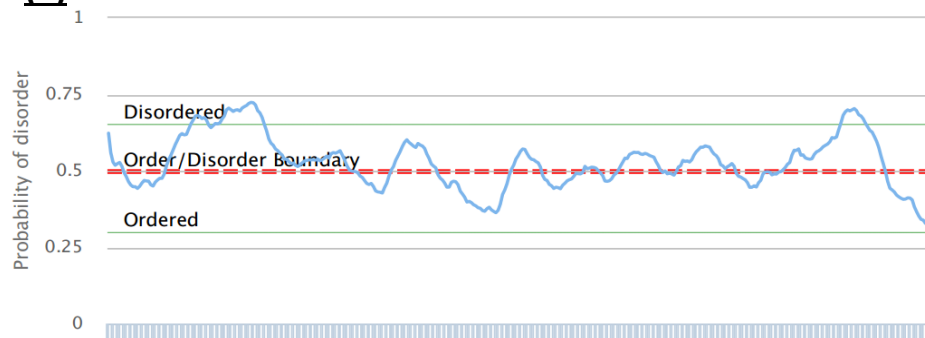


Coil

**(B)**



**(C)**



**(D)**

```
TbCAP51      MSTTALSSAKNARSPGSNAAVSPGSKLASLREGASTGVAKDEAESALGGLAGGDA----
TbCAP51V     MSTTALSSAKNARSPGSNAAVSPGSKSVSLREGALRSDRKSTRELSPQSELKSDRKSMRE
               ***** .***** . * . * : . * ..

TbCAP51      -----EGLAAGAMSGLAETATEA----AADELGDAASEDQANFF
TbCAP51V     ISHSELKSDRKSTRELSPQSELKSDRKSMGVSPQSELKSDRKSMREISHSELK--S--D
               :***:: : : * :..: : :

TbCAP51      SEKLVEVIPQCN-----EIEGERD-LNGELG--
TbCAP51V     RKSTRELSPQSELKSDRKSVRDSEMRSAKSGRSVEARYGVRDDGDGEWSELGSEVTSEL
               :. * : **.: : ** . *.*: .

TbCAP51      -----EEAPRTDEGDVNNVG-----CS----SAAESLPQEEVEETEEQI
TbCAP51V     RSKAKSGRRSERRQGGADEFELSESVDVAHGENFFDSSRRSDRQPSSGAPEEEEEETEEQI
               : . .*.**: * . * . * ** *****

TbCAP51      IIRLQKELREFEAYEKEMQPNLNIYEDIIELRREAEQLEDDRKRALLQIDLGDISSYYDG
TbCAP51V     IIRLREELREFEAYEEEMQPNLNMYEEVIELRREAEQLEDDRKRALLQIDLGDISSYYDG
               ***: :*****:*****:*: :*****:*****:*****

TbCAP51      DVEREVEELEAAVENSYLQACWTKERQRYIARNETDLPEWWRVNVKDLPSRLKDARASFD
TbCAP51V     DVEREVEELEAAVENSYLQACWTKERQRYIARNETDLPEWWRVDIMKLGFRLLREAKRSFD
               *****: . * **:*: : **

TbCAP51      SSNRVAQDLYRAHEVKVNATTDSDRKAKAEIFNGMKAIEATKGERERLRQLQKDQLFHL
TbCAP51V     SSNRVAQDLYRAHEVKVNATTDSDRKAKAELRQEIAAELNELRQRERLLVIDKEQRFHL
               *****: : **: : : .*** :*: * **

TbCAP51      RRGTHVKDQLLTITTKERDIREARVADMESKLVQQRYSALNNEMKELNGDIEGLKRLLTDQK
```

TbCAP51V	RRGTHVKELQVTSKRRQDALEARVVGEEENRHFGASVELKRTVKDLEMELIQLRRELKDAK *****: :*:.:*: * * * * . * . : . * : . : * : * : : : * : * * . * *
TbCAP51	HASRETLETLLKKQHVAVDSSRGELSEAREKEYERENSELMLLKHDLTQTVLHYIRVRAREAD
TbCAP51V	VSGESIVQELRKHLLDVTRSGEEVRKARQDYEEENS DLM SLKHELQCVLHYIRVRAREAN : . . . : : * : * : : * * * : : * : . * : . * : * : * : * : * : * : * : * : * : * : * : * : * : *
TbCAP51	K
TbCAP51V	Y

## (E)

L.major	MSAQQQPSE-----ENSDRMVAARTKAEAAARMAVERRRSMNSDAQERSNSPDHHDTPPTPL
TbCAP5	MSTTALSSAKNARSPGSNAAVSPGSK-----LASL-REGASTGVAK---DEAESAAALGGL
T.cruzi	-----METNVET---QRQESGSPKEL : . . . . : . . : *
L.major	RGATQERSSS---VDLANGGI---EDPHGDSGEHR-PNSSIVKRRSSMPKNGVRDGSRRG
TbCAP5	AGGDAEGLAAGAMSGLAETATEAAADELGDAASEDQANFFSEKLVEVIPQCNEIEGERDL
T.cruzi	AGTRA-----QKAPEAAAEFDYRELGEAAVQEFEE----- * : . * : . . :
L.major	SGVV-AVGP-----PNASYSGTKKTADPQ---TPEEYAAAIEEMRQEYERIIMKDAE
TbCAP5	NGELLGEEAPRTDEGDVNNVGCSSAAE-SLPQEEV-EETEEQIIIRLQKELREFEAYEKE
T.cruzi	-----YSSASG-AAPGRRSSIRFVLASIEEKENELRELEAREKA * . : : * . * . : * . : :
L.major	LEERLQLYSDVVGLRKELVIVQEDEARLRQQLAATEAVIASSDPTVKKVPIYEEEAQQS
TbCAP5	MQPNLNIYEDIIELRREAEQLEDDRKRALLQIDL-GDISSYYDGDVEREVEELEAAVENS
T.cruzi	VESQLQVYSEVIGLRKELRLVKEETERVRRQLDSINVFAEKYGNABERDLNSAEAAQRS : : . * : * : : * : * : : : * * : . . * : : * . : *
L.major	VLQKLWMESEEC-----HNPDTMEWAAIDVNNLRERVDAAHKKFTAAEMADSLYLQEEEA
TbCAP5	YLQACWTKERQRYIARNETDLPEWWRVNVKDLPSRLKDARASFDSSNRVAQDLYRAHEVK
T.cruzi	HVQQVWTSVCRERPNESSLGDPSSWTKVDLDNLSRRVEAARGKLQAASNAAKVYASQEDM : * * . . * . * : : : * * : . * : : : . * . : * :
L.major	INKTTDARERERATIAENHEREMEGELSEARVHARQVAAEQHFHRHRGTAQAPPVVLTDQDK
TbCAP5	VNATDTSRDKAKAEIFNGMKAEIEATKGERERLRQLQKDLFHLRRGTHVKDQLLTIT-K
T.cruzi	INKNTEERDKAKAKLLESVGVEMAKLVETRSLKQLDDQRYHFRRGTHVKQRHVPSL-K : * . * : * : : * : . * : : * : : * : * : * : * : * : *
L.major	KRTTRQRCVDEVEFHTTAQVGRMKDELTELMEQVKMLKRHLDDSRQVTEKKRREYEATLK
TbCAP5	ERDIAREARVADMESKLVQQRYALNNEMKELNGDIEGLKRLLTQKHASRETLETLLKKQHV
T.cruzi	EQDIMKGRVYNDESQRQFGDVLLDECKDLTEWVNGAKRELEDNKRYSYAKEQSELEEQLQ : : : * : * : : : : * . * : : * * * * : : : :
L.major	AVEADGSEAREMKEGLLKEEEELKELKADIQAVLHYVRAKNREEEGW
TbCAP5	AVDSSRGELSEAREKEYERENSELMLLKHDLTQTVLHYIRVRAREADK-
T.cruzi	VMEDNGKEARELRAFFERENDQLKGLKQDLQHVLIYVRMRNREAF-T- : : . * * : : : * : * * : * * * : * : *

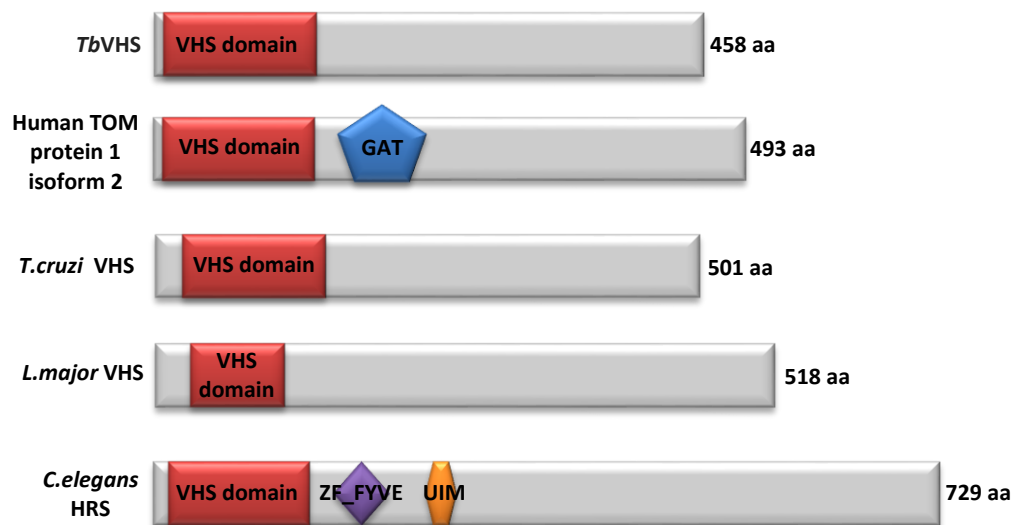
**Figure 4.8 (A)** Interpro Scan analysis of *TbCAP51*, no domain structures were detected. However, results shows possible coil coil within the sequence. **(B)** Disorder prediction of structured regions performed by IUPRED. **(C)** Disorder prediction of structured regions performed by RONN. **(D)** Multiple amino acid sequence alignments between *TbCAP51* (*Tb927.7.2640*) against it paralouges *TbCAP51V* (*Tb927.7.2650*). **(E)** Multiple amino acid sequence alignments between *TbCAP51*, top hit in *T.cruzi* (TcCLB.506859.180) and *L.major* top hit (LMJLV39\_220012800) by Clustal Omega using default parameters. “\*” indicates identical aa and “:” indicates conserved.

## 4.8 VHS (VPS-27, Hrs and STAM)-domain containing protein

The VHS domain containing protein (*Tb*927.11.10770) was identified by BioID/SILAC when whole cell samples, rather than cytoskeletons, were used for the analysis. In almost all cases, the VHS domain can be found only at the N-terminus of proteins and the roles of these proteins are mainly associated with vesicular/endocytosis trafficking (Baiady et al., 2016). In addition, VHS domain-containing proteins often localise to membranes through interactions with the membrane and/or other mechanism like endocytosis (Misra et al., 2000). Moreover, evidence also suggests that VHS is also linked to other domains such as zinc finger-FYVE, SH3 and TAM, and together they play important functions in membrane binding (Mao et al., 2000). I used this *Tb*927.11.10770 sequence to perform BLAST analysis against the restricted set of eukaryotic organisms shown in Table 4.1. This NCBI BLAST analysis showed that proteins with  $e^{-10}$  value were only found in humans, *C. elegans* and trypanosomatids. In humans, a gene identified as a TOM (Target Of the Myb) protein 1 isoform 2 came back as the top hit (domain structures are shown in Fig 4.9-A). A reciprocal BLAST confirmed that *Tb*VHS (*Tb*927.11.10770) was the top hit of the human TOM protein 1 isoform 2. This human protein (NP\_001129204.1) shares N-terminal VHS domain with *T. brucei*, and also consists of a domain called a GAT domain, which can be found in other vertebrate TOM proteins and eukaryotic GGAs (for Golgi-localized, gamma ear-containing ADP ribosylation factor (ARF)-binding proteins). The function of the GAT domain is involved in trans-Golgi network membranes, and interaction with small GTPases (GTP-bound form); additionally, it can also bind specifically to the Rab GTPase effector rabaptin5 and ubiquitin (Collins et al., 2003; Shiba et al., 2003; Yamakami et al., 2003; Zhu et al., 2004). Based on InterproScan/Prosit analysis, the VHS domain of both *T. cruzi* (TcCLB.458015.10) and *L. major* (LmjF.36.5140) proteins are all located at N-terminus just like in *T. brucei*. In *C. elegans*, a HRS (Hepatocyte growth factor-Regulated TK Substrate) protein (NP\_501375.2) came back as the top hit to *Tb*VHS. The domain structure analysis demonstrated that in addition to a VHS domain at the N-terminus, there is also a Zinc finger FYVE/FYVE-related domain and Ubiquitin-

interacting motif (UIM) domain in the *C. elegans* protein. As the last step of the *Tb*VHS analysis, I used CLUSTAL OMEGA to align the *Tb*VHS sequence with human TOM protein 1 isoform 2, and the *T. cruzi* and *L. major* VHS proteins, as well as the *C. elegans* HRS protein sequence (Fig 4.9-B). The results show good conservation in the VHS domains of all five proteins (yellow highlights), other domains, which does not appear in the *T. brucei* sequence are coloured in blue, purple and green.

**(A)**



[illegible]



TbVHS	SS---EEV-----SEERNRKPVKH-----
T.cruzi	TK---ADMPPSEPIPSVPEKEEEVLMAGKEENQSKQEETQKQDPDPHE-----KS
C.elegans	--KHAMLMEQREQALQRFQQQQEMAMRRHQQQAYYNPQQMGYGAPPPSGPQQPPQQSY
H.sapiens	EEPKG---VTSEGKFDKFLEER---AKAADRLP-----NLSSPSAEGPPGPPSG---
L.major	PTPPAADLPKDVDVDFDTFLEGH---GVS-----
TbVHS	-----VYTDFEDLLARQ---APK-----
T.cruzi	PEEKSKQPAGEDDEFDAFLEGR---LSK-----
	:: :   :
C.elegans	GYQQGPQSQAPPPSSQYQQSHASTQQQQQQYYQHYQANQTVTQPVQQAQQQYQQYQGY
H.sapiens	-----P-----APRKKTQE-----KDDDMLFAL-----
L.major	-----
TbVHS	-----
T.cruzi	-----
C.elegans	QQQQPYQQQAGYQNPQQHQNYQNGTTNENGQYQQQSSSEVKQEHQNHGYQQQQAQHPPAS
H.sapiens	-----
L.major	-----
TbVHS	-----
T.cruzi	-----
C.elegans	VASNGHNGYGNVDQNAHHHQQQQPQIAEQPLISFD
H.sapiens	-----
L.major	-----
TbVHS	-----
T.cruzi	-----

**Figure 4.9 (A)** Schematic cartoons of VHS domain structure within different organisms. Different domains are indicated by different colours. The VHS domain is normally located at N-terminus of the protein sequences (yellow highlights); In human TOM protein 1 isoform 2 also contains GAT domain (blue highlights); The Zinc finger FYVE/FYVE-related domain (purple highlights) and Ubiquitin-interacting motif (UIM) domain (green highlights) in *C. elegans*. **(B)** Multiple sequence alignments between *TbVHS* protein against human, *T. cruzi*, *L. major* and *C. elegans* by Clustal Omega analysis. “\*” indicates identical aa and “:” indicates conserved aa.

## 4.9 P25- $\alpha$ /TPPP (Tubulin-Polymerization-Promoting Protein)

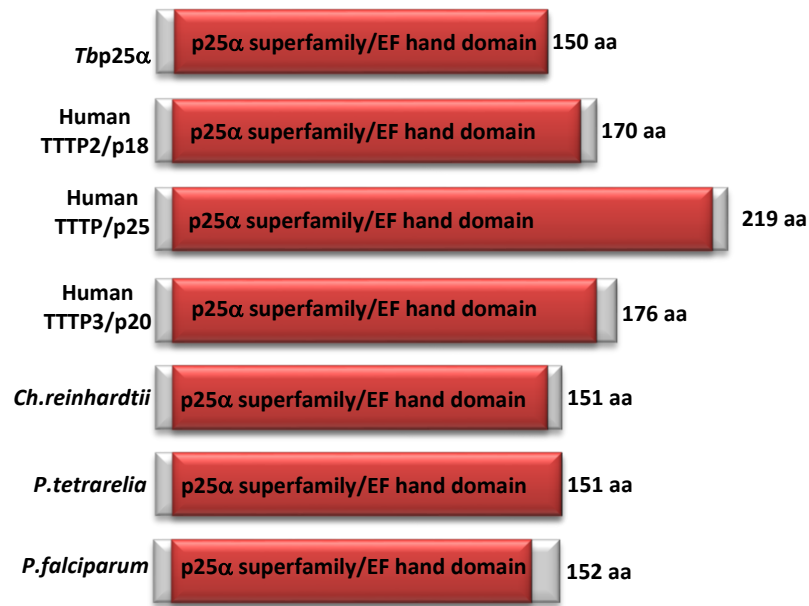
The TPPP or P25- $\alpha$  is a 25 kDa protein originally identified as a brain specific protein that could be phosphorylated by Ser/Thr-Pro kinase (TPK II) (Takahashi et al., 1991). In eukaryotic cells, TPPP/p25 co-localises selectively with microtubules and promotes tubulin polymerization into normal and double-walled microtubules and induces their bundling activities. The role of TPPP/p25 is thought to be similar to Microtubule Associated Protein (MAP)-like stabilization of the microtubular network (DeBonis et al., 2015). Overexpression of TPPP/p25 can lead to protein aggregation in transfected HeLa cells, this mimics the characteristic aggresome-like formations found in Lewy bodies (which contains  $\alpha$ -synuclein, ubiquitin and neurofilaments (Lehotzky et al., 2004), which enriched in human brains in certain neuropathological conditions. For instance, in neurodegenerative disorders such as Parkinson's disease and other synucleinopathies, TPPP/p25 shows complete co-localisation with  $\alpha$ -synuclein in Lewy bodies (Kovács et al., 2004; Lindersson et al., 2005; Orosz et al., 2004). Later studies reported that the TPPP/p25 gene is conserved in organisms that produces cilia/flagella, but is absent from those that are non-ciliated organisms, this may suggest TPPP/p25 is associated with basal bodies/centrosomes and/or flagella structure (Orosz and Ovádi, 2008; Ovádi et al., 2005).

In the *T. brucei* genome, there is only one TPPP-like protein (*Tb927.4.2740*). An initial BLAST search revealed that all three non-ciliated organisms failed to encode a TPPP-like gene in their genomes. In the selected flagellated organisms, only *G. intestinalis* and *C. elegans* failed this NCBI protein BLAST. However, it should be borne in mind that here I used the *T. brucei* sequence as my query and not the human TPPP protein. The top hit human protein came back as another TPPP-like protein superfamily member called Tubulin polymerisation-promoting protein family member 2 (TPPP2, also known as p18). The actual function of TPPP2 is not understood, but data suggests that p18 may also be involved in tubulin binding but has no microtubule bundling activity (Vincze et al., 2006). In the human genome, one more gene sequence shows

homology with TPPP/p25 and TPPP2/p18; this protein is termed as p20 or TPPP3. These three TPPP-like proteins are identified in divergent eukaryotes that share a central p25 domain, and they display high amino acid sequence homology, apart from the N-terminal tail of TPPP/p25 (see Fig 4.10-C) (Vincze et al., 2006). In addition, the data also indicates that while all three proteins are highly conserved in amino acid sequences, their structures are very different. TPPP/p25 has low helical content, but is more flexible and dynamic like other MAPs (Dunker et al., 2002; Orosz et al., 2004; Uversky, 2002). TPPP3/p20 shows similar characterisation to TPPP/p25 and behaves more like a disordered protein, while TPPP2/p18 has a rather stable structure and behaves in a more ordered fashion (Vincze et al., 2006). The reasons why the TPPP-like proteins are predicted to behave differently is unclear. However, it cannot simply explained by N-terminal region differences (Otzen et al., 2005; Zotter et al., 2011).

Domain structures of TPPP-like proteins in a variety of organisms are shown in Fig 4.10-A. When I used InterPro Scan to analyse the domain architectures for each protein, the program also identified an EF-hand domain (participation in  $\text{Ca}^{2+}$ -binding) shown on the cartoon (Fig 4.10-A). The CLUSTAL OMEGA sequences alignment between *Tbp25 $\alpha$*  and human, *T. brucei*, *L. major*, *P. falciparum*, *C. reinhardtii* as well as *P. tetraurelia*, are shown in Fig 4.10-B to F respectively. The results show good sequence conservation within the central p25 domains of these proteins (yellow highlights).

## (A)



## (B)

TPPP2/p18 MASEAEKT FHRFAAFGESSSSGTEMNNKNFSKLCKDCGIMDGKTVTSTDVDIVFSKVKAK  
Tbp25a ----MEAVFYAFASFGTAP--TKEMDNAHFSKMLKEAKI-IGKTFSTADLLFNKIKAK  
\* .\*: \*\*:\*\*\* : .\*: \* :\*\*\*: \*:. \* \*\*\*.\*\*\*\*.\*\*\*:\*.\*\*\*

TPPP2/p18 NARTITFQQFKEAVKELGQKRFGKSPDEVLENIYGLMEGKDPATTGATKATTVGAVDRL  
Tbp25a GARKITTFEFNTRALPDIATKLK-MTPEQVAE----ILTKASPASNS-T---KAEAVKFH  
.\*\*\*.\*\*\* :\*: . :\*: :\*\*\*: \* :\*: .\*\*\*. \* ..\*\*.

TPPP2/p18 TDTSKYTGTGTHKERFDESGKGKGIAGREEMTDNTGYVSGYKSGSTYDKKTK-----  
Tbp25a DDKNLYTGVIYKAG-----GPTNVDRNAGSLSGVVDRRVDQVDVRGTTSSQK  
\*..\*\*\*.\* \* :\*: \* :\*\*\* :. :.:

## (C)

TPPP2/p18 -----MASEAEKT FHRFAA  
TPPP/p25 MADKAKPAKAANRTPPKSPGDPSKDRAAKRLSLESEGAGEGAAASPELSALEEAFRRFAV  
TPPP3/p20 -----MAASTDMAGLEESFRKFAI  
: \* : \* : \*

TPPP2/p18 FGESSSSGTEMNNKNFSKLCKDCGIMDGKTVTSTDVDIVFSKVKAKNARTITFQQFKEAV  
TPPP/p25 HGDARATGREMHGKNWSKLCKDCQVIDGRNVTVDVDIVFSKIKGKSCRTITFEQFQEAL  
TPPP3/p20 HGDPKASGQEMNGKNWAKLCKDCKVADGKSVTGTVDIVFSKVKGKSARVINYEYEFKKAL  
. \* : :\*: \*\* : \* : \* : \* : \* : \* : \* : \* : \* : \* : \* : \* : \*

TPPP2/p18 KELGQKRFGKSPDEVLENIYGLMEGKDPATTGATKATTVGAVDRLTDTSKYTGTGTHKERF  
TPPP/p25 EELAKKRFDKSSSEAVREVRHLIEGKAPIISGVTKAISSPTVSRLTDTTKFTGSHKERF  
TPPP3/p20 EELATKRFGKSKEEAFDAICQLVAGKEPANVGVTAKTGGAVDRLTDSRYTGSHKERF  
\*: \* : \* : \* : \* : \* : \* : \* : \* : \* : \* : \* : \* : \* : \*

TPPP2/p18 DESGKGKGIAGREEMTDNTGYVSGYKSGSTYDKKTK---  
TPPP/p25 DPSGKGKAGRVLDVDESGYVSGYKHAGTYDQKVQGGK  
TPPP3/p20 DESGKGKGIAGRQDILDDSGYVSAYKNAGTYDAKVKK--  
\* \* : \* : \* : \* : \* : \* : \*

## (D)

```
TbP25a      --MEAVFYAFASFGTAPTKEMDNAHFSKMLKEAKIIGKTFSTADLLFNKIKAKGARKI
TcCLB.504199.20 MSIESAFYAFASFGGAPTKEMDNAHFSKMLKETKVIGKQFTSTDADLLFNKVKAKGARKI
TcCLB.506635.130 MSIESAFYAFASFGGAPTKEMDNAHFSKMLKETKVIGKQFTSTDADLLFNKVKAKGARKI
               :*:.***** *****:*.*** *****:*****

TbP25a      TFTEFNTRALPDIATKLMKMTPEQVAEILTKASPASNSTKAEAVKFHDDKNLYTGVYKAGG
TcCLB.504199.20 TLSDFVDKAVPEIASKLKSAEELIADISSCSPEARATKADAVKFHDDKNMYTGVYKAGG
TcCLB.506635.130 TLSDFVDKAVPEIASKLKSAEELIADISSCSPEARATKADAVKFHDDKNMYTGVYKAGG
               *:.* :*:*:*:*:* :*:* :*..** :*:*:*:*:*:*:*:*:*:*

TbP25a      PTNVDRNAGSLSGVVDRRVDQVDVRGTTSSQK
TcCLB.504199.20 PTNVDRNAGSLSGVVDRRVAQTDVRGTTASQK
TcCLB.506635.130 PTNVDRNAGSLSGVVDRRVAQTDVCGTTASQK
               *****:***** *.* ***:***
```

## (E)

```
TbP25a      ---MEAVFYAFASFGTAPTKEMDNAHFSKMLKEAKIIGKTFSTADLLFNKIKAKGARK
LMJSD75_340023800 -----MDNSHFSKMLKECKIIGKSFTSTDADLLFSKVKAKEARK
LmjF.34_1520 MDNFQATFEAFASFGSAPSKEMDNSHFSKMLKECKIIGKSFTSTDADLLFSKVKAKEARK
LmjF.34.1530 -----MDNSHFSKMLKECKIIGKSFTSTDADLLFSKVKAKEARK
               ***:*****.*****:*****.*:*** **

TbP25a      ITFTEFNTRALPDIATKLMKMTPEQVAEILTK-----
LMJSD75_340023800 ISFTEFKEKAIPeIAAKMKKTPADIEAMITNAAPKSSGKTADTVRFHDDKSTYTGAAKQG
LmjF.34_1520 ISFTEFKEKAIPeIAAKMKKTPADIEAMIA-----
LmjF.34.1530 ISFTEFKEKAIPeIAAKMKKTPADIEAMIA-----
               *:***: :*:*:*:*:* ** : : :

TbP25a      -----ASPASNSTKAEAVKFHDDKNL
LMJSD75_340023800 ADERGPKRRLARGCRGKAIPeIAAKMKKTPADIEAMIANAAPKSSGKTADTVRFHDDKST
LmjF.34_1520 -----NAAPKSSGKTADTVRFHDDKST
LmjF.34.1530 -----NAAPKSSGKTADTVRFHDDKST
               *: * ..***:*.****.

TbP25a      YTGvyKAGGPTNVDRNAGSLSGVVDRRVDQVDVRGTTSSQK
LMJSD75_340023800 YTGAAKQGGPTNVDRNAGSLAGVVDRRQETINNRGFTA---
LmjF.34_1520 YTGAAKQGGPTNVDRNAGSLAGVVDRRQETIDNRGTTAKQI
LmjF.34.1530 YTGAAKQGGPTNVDRNAGSLAGVVDRRQETINNRGFTA---
               ***. * *****:***** : : * * *
```

## (F)

TPPP/p18	MASEAEKT	FHRFAAFGESSSSGTEMNNKNSKLCCKDCGIMDGKTVTSTDVDIVFSKVKA
P.vivax	----MENAFYIYTK-----	NEADMDSRTFVKILKDAKLLS-KKLTAVDADLTFKVKAK
P.tetraurelia	MQGNVQQVF	LQFTA-----NKPEMDGKTFKAVSKDCHLLD-KKLTSTDVDLIFAKIKPT
Ch.reihardtii	MSDALKNAF	IAFASYGKGQMMKQMDNKNFSKCIKDSGILD-KVITSTEVDITFMKVKA
LmjF.34.1520	-MDNFQATFEAFASFGSA--	PSKEMDNSHFSKMLKECKIIG-KSFTSTDADLLFSKVKA
Tbp25a	----MEAVFYAFASFGTA--	PTKEMDNAHFSKMLKEAKIIG-KTFTSTDADLLFNKIKAK
TcCLB.504199.20	--MSIESAFYAFASFGGA--	PTKEMDNAHFSKMLKETKVIG-KQFTSTDADLLFNKVKA
	: . * ::	: * . * * * : : . * . * . * . * : * * * .
TPPP/p18	N-ARTITFQQFKE-AVKELGQKRFGKSPDEVLENIYGLMEGKDPATTGATKATTVGA	VD
P.vivax	G-AKRINYDQFVE-AVKHLVDKH--	KLDYDQFVEKLCNE-ASSGPI----LYGKAANVR
P.tetraurelia	PAARSITYAQFEK-GLQMMAEKK--	GVGVQDVHNQIL---NAGGPH---FQGTKADAVK
Ch.reihardtii	T-DRTINFAQFCT-ALEHFAAKR--	GVSVDSLHAK-V---EAASPT----SNATQAEAVK
LmjF.34.1520	E-ARKISFTEFKEKAIPETIAAKM--	KKTPADIEAM-I---ANAAPK----SSGTKADTVR
Tbp25a	G-ARKITFTEFNTRALPDIA TKL--	KMTPEQVAEI-L---TKASPA----SNSTKAEAVK
TcCLB.504199.20	G-ARKITLSDFVDKAVPEIASKL--	KKSAEELIAD-I---SSCSPE----ARATKADAVK
	: * . : * . : : *	. . * . * . *
TPPP/p18	RLTDTSKYTGTHERFDES-----	GKGKGIAGREEM-TDNTGYVSGYKSGTYDKKT
P.vivax	FHDDKSTYTG VHKMG GPTTV DKNKTHFSDISEITDRSEC	NIRGVN-----LSVEKN
P.tetraurelia	FHDDKNLYTGVHANGGPSTIDKN-H--	GGLNTICDRSQA--DVRGSQ-----KMMKNI
Ch.reihardtii	FHDDKNLYTGVYKNGGPTNIDKQAA--	GGLAGHLDRSPA--DVRGVK-----F-----
LmjF.34.1520	FHDDKSTYTGAAKQGGPTNVDRN-A--	GSLAGVDRRQETIDNRGTT-----AKQI--
Tbp25a	FHDDKNLYTGVYKAGGPTNVDRN-A--	GSLSGVDRRVDQVDVRGTT-----SSQK--
TcCLB.504199.20	FHDDKNMYTGVYKAGGPTNVDRN-S--	GSLSGVDRRVAQTDVRGTT-----ASQK--
	* . . * * . .	. . * : *
TPPP/p18	K	
P.vivax	L	
P.tetraurelia	-	
Ch.reihardtii	-	
LmjF.34.1520	-	
Tbp25a	-	
TcCLB.504199.20	-	

**Figure 4.10 (A)** Schematic cartoons of p25 domain structure within different organisms. **(B, D-E)** Multiple sequence alignments between *Tbp25α* protein against human, *T. cruzi*, *L. major*, *P. falciparum*, *C. reinhardtii*, and *P. tetraurelia* using the ClustalOmega analysis. **(C)** Multiple sequence alignments between three human TPPP-like proteins. **(F)** Multiple sequence alignments compared all TPPP-like proteins that came back with good e-value(s) when *Tbp25α* was the query sequences. The p25 domain/EF hand domain is normally located at central of the protein sequences (yellow highlights) “\*” indicates identical aa and “:” indicates conserved aa.

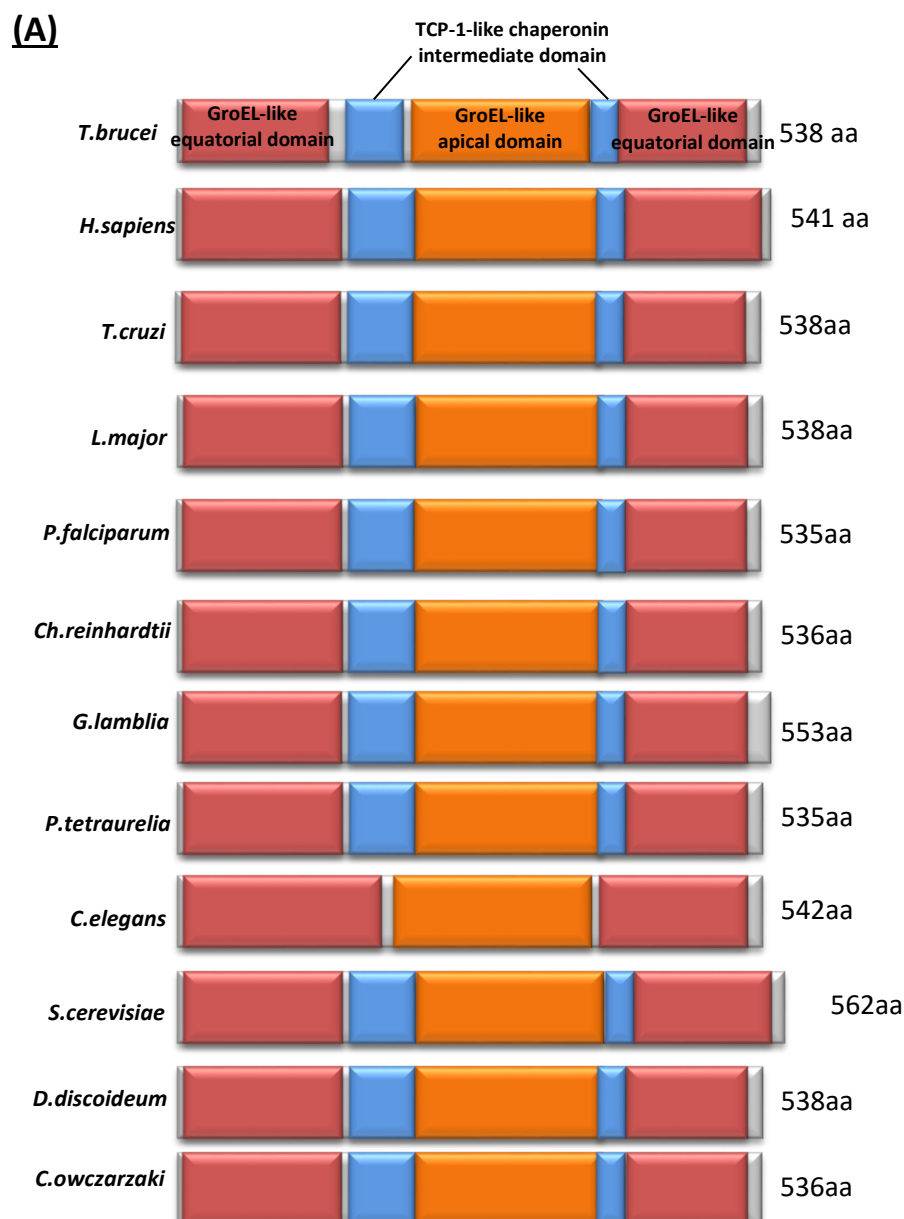
## 4.10 TCP-1-ε (T-Complex Protein 1 subunit epsilon)

The TCP-1-ε is a molecular chaperone that belongs to chaperonin containing TCP-1 complex (CCT), also known as the TCP1 ring complex (TRiC). The common structure of this complex consists of two identical stacked rings, each containing eight different proteins, but homologous subunits, which are designated as CCTα, CCTβ, CCTγ, CCTδ, CCTε, CCTζ, CCTη, and CCTθ (Liou and Willison, 1997). TCP-1 complex was first identified as chaperone for actin and tubulin filaments assembly in biogenesis (Yaffe et al., 1992), later studies demonstrated that TCP-1 also involved in the folding of newly synthesised cytosolic protein substrates (Dekker et al., 2008; Spiess et al., 2004; Yam et al., 2008). Furthermore, proteomic and genomic approaches reveal that TCP-1 complex interacts with a large number of proteins associated with either the spindle pore, protein degradation pathways or the cell wall/membrane (Dekker et al., 2008). Like all the chaperones, TCP-1 complex can bind with ATP, but rather through cofactors, it has a built-in lid located in the apical domain of each subunit responsible for closing the central cavity (Leitner et al., 2012).

In human, mutation in TCP-1-ε/CCT5 gene can lead to hereditary sensory and autonomic neuropathy with spastic paraplegia (HSNSP) (Min et al., 2014). Moreover, evidence also suggests that as part of the BBS/CCT complex, it might also be involved in BBSome assembly in order to regulate the transport of vesicles to cilia during ciliogenesis (Seo et al., 2010). The TCP-1 complex is conserved in all organisms; eukaryotes, archaeobacteria and eubacteria (Kubota et al., 1995).

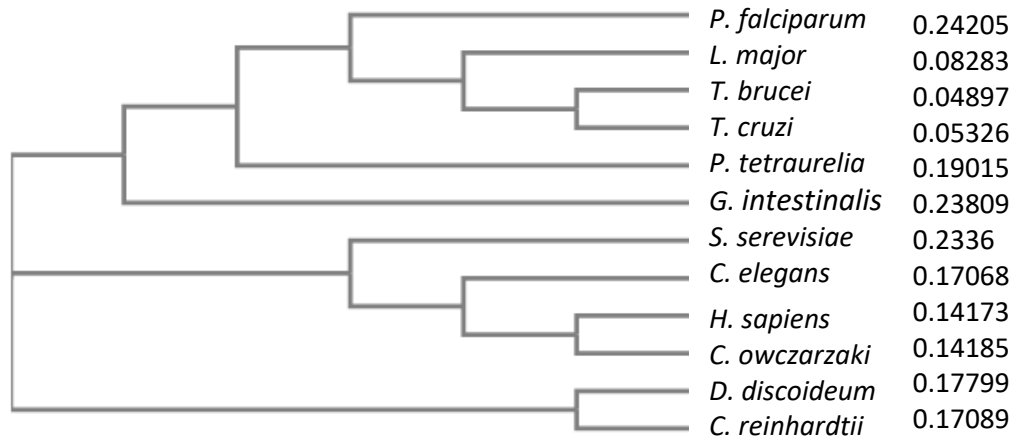
At the starting point for my bioinformatics analysis, I performed a NCBI BLAST search, using *Tb*TCP-1-ε as my query against human and all the organisms listed in the Table 4.1. Considering the critical roles of TCP-1 complex in protein folding it was not surprising that all the selected organisms appear to have multiple TCP-1 genes. Because there are so many hits with a good e value, I only took the top hit from each organism for further domain and multiple sequence alignment. Fig 4.11 –A shows the domain structures of all organisms in the table. Both ciliated and non-ciliated

organisms share extremely similar domain architectures. In most of the organisms (apart from in *C. elegans*), three possible domains are detected by InterproScan analysis; (1) GroEL-like equatorial domain that carries ATPase activity (2) GroEL-like apical domain, and (3) TCP-1-like chaperonin intermediate domain involved in substrate binding (Pappenberger et al., 2002; Roseman et al., 1996). The CLUSTAL OMEGA sequences alignment between *Tb*TCP-1-ε to the top hit TCP-1 protein from each organism is shown in Fig 4.11-B. The results reveal good conservation within the different domains of the proteins (sequence alignments see Appendix II Fig 9.8).





**(B)**



**Figure 4.11 (A)** Schematic cartoons shows the domain structures of TCP-1- $\epsilon$  in different organisms. Different domains are indicated by different colours. (1) GroEL-like equatorial domains that locate at both C and N terminus are in red; (2) GroEL-like apical domain is located at the central of the sequences in orange; (3) TCP-1-like chaperonin intermediate domain, which can be found between the GroEL-like equatorial and apical domain shows in blue. **(B)** Phylogenetic analysis of the *Tb*TCP-1- $\epsilon$  proteins by CLUSTAL OMEGA.

## 4.11 Protein Kinase (*Tb927.9.6560*)

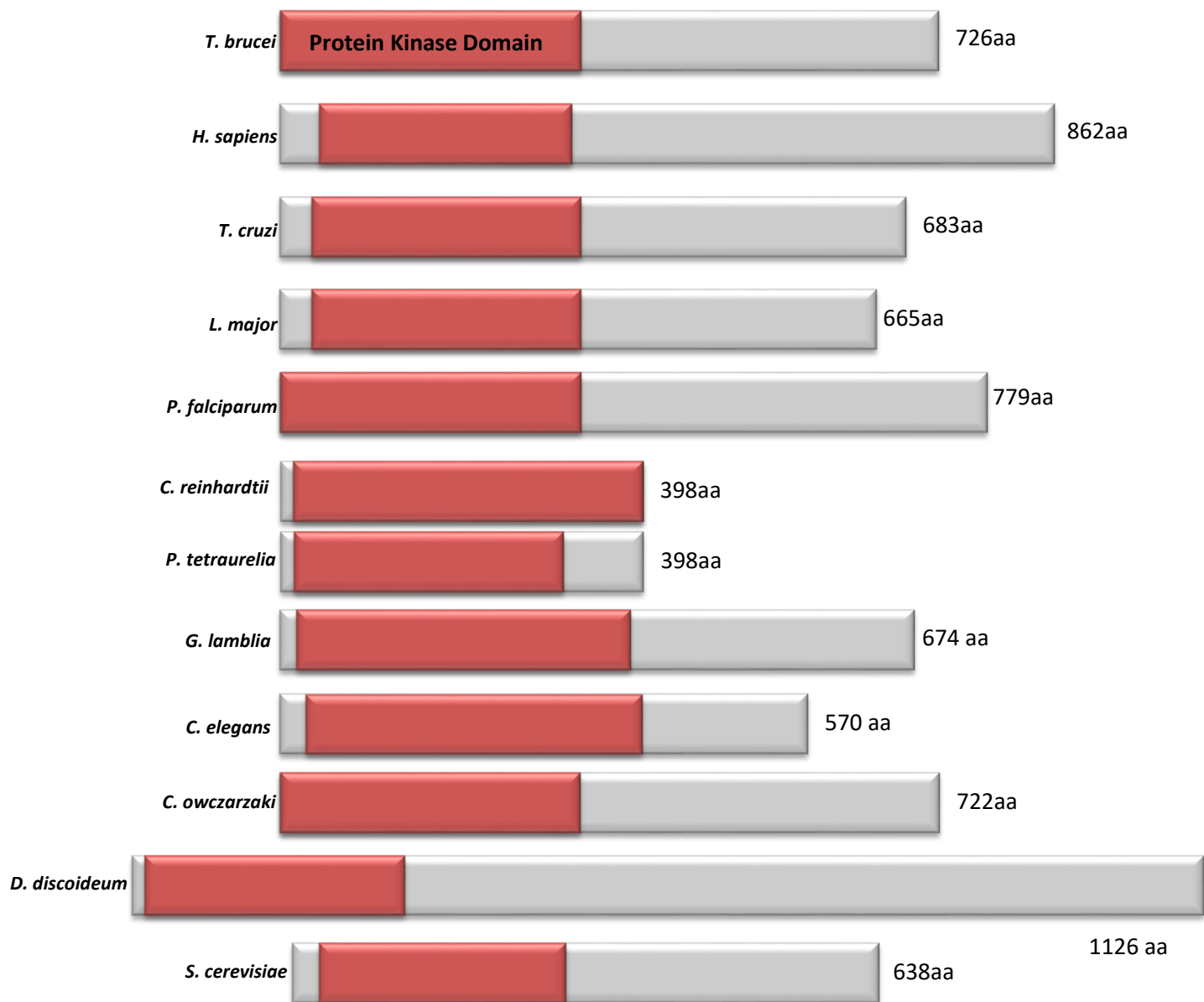
In eukaryotes, protein kinases are enzymes that play an essential role in communicating intracellular signals within cells and are involved in many cellular processes, such as cell division, proliferation, apoptosis, and differentiation (Manning et al., 2002a). In humans, there are more than 568 members within the eukaryote protein kinase (ePKs) family, which constitutes more than 2% of the entire human genome (Manning et al., 2002b). ePKs share a high degree of conservation at the amino acid sequence level and in their catalytic domain structures (Hanks et al., 1988; Taylor and Radzio-Andzelm, 1994). Among all the ePKs, the majority of members are involved in phosphorylation of proteins on serine/threonine residues, and are derived from protein-serine/threonine kinases (STKs) that evolved in the early metazoan evolution (Rokas et al., 2005). There is another group of kinases, named as atypical protein kinases. Family members of this kinase group share less sequence identity and structural similarity to most of the typical protein kinases. There is less information on these atypical protein kinases, however, some studies suggest they diverged early in evolution and have followed a distinct evolutionary track (Leonard et al., 1998; Middelbeek et al., 2010). Protein kinases are also found in bacteria, plants, worms and flies. However, some of the conserved functions or pathways have been lost or expanded during each evolution lineage (Manning et al., 2002a). For example, some kinases can be found in both humans and flies and share conservation in morphogenesis and neurobiology function, but are missing from the worm genome, which indicates these functions could have evolved after nematodes diverged from the main metazoan lineage (Manning et al., 2002a).

The protein kinase encoded by *Tb927.9.6560* is 726 amino acid long; the protein kinase domain covers the N-terminus and nearly reaches the central region. The *Tb927.9.6560* protein kinase is currently annotated as a NAK pseudokinase (Parsons et al., 2005), as it is missing an essential catalytic residue. When I performed a NCBI BLAST using the *Tb927.9.6560* as my query, all of the chosen organisms came back with good e values. Again, for further protein domain and multiple sequence alignment I only took the top hit for each organism. The protein domain analysis was

performed by InterProScan/Prosit (Fig 4.12-A), and all proteins encoded a protein kinase domain within their amino acid sequence. However, where the domain is located differed slightly from organism to organism, and the length of the top hit proteins also varied. Proteins from most of the organisms shown in Fig 4.12 appear to have N-terminal protein kinase domain; apart from *C. reinhardtii* where the protein kinase domain covers nearly the entire sequence. Some data has suggested that the N-terminal protein kinase domain is normally a glycine-rich region close to a lysine residue, which has been shown to be involved in ATP binding. When the protein kinase domain is located in the central part, it consists a conserved aspartic acid residue which plays an important role in the catalytic activity of the enzyme (Knighton et al., 1991). *Tb*protein kinase (*Tb*927.9.6560) only shares 30% amino acid identity to human Adaptor-Associated Kinase 1 (AAK1) protein (AAI43711.1), which acts as a positive regulator in the Notch pathway (Gupta-Rossi et al., 2011). To date, in *T. brucei* there are 158 eukaryotic PKs (ePKs), 12 predicted pseudokinases and 20 atypical PKs (aPKs; despite their lack similarly of conserved sequence to the ePK domain HMM, but aPKs still have protein kinase activity) (Nett et al., 2009; Parsons et al., 2005; Urbaniak et al., 2012).

I have also performed a reciprocal BLAST using the human AAK1 protein as the query sequence to check if there are any other protein kinases in *T. brucei* that are more similar to this human protein. This analysis showed that the Protein Kinase (*Tb*927.9.6560) was the second hit, and a protein called NIMA-related protein kinase (*Tb*927.10.460) came back as the top hit. Pradel et al., (2006) have shown that this protein is a basal body-cytoskeletal NIMA-related kinase (NRK) protein (*Tb*NEKC), which plays critical roles in basal body segregation. RNAi knockdown of *Tb*NRKC and overexpression of kinase-null *Tb*NRKC can lead to the accumulation of cells with four basal bodies and cytokinesis defects (Pradel et al., 2006). CLUSTAL OMEGA sequence alignment comparisons between *Tb*protein kinase and the top hit of protein kinase in each organism is shown in Fig 4.12-B. The results reveal that when all sequences were compared they did not show good amino acid conservation. However, when I just used *T. brucei*, *T. cruzi* and *L. major* the conservation at the N-terminus was much improved Fig 4.12-C.

**(A)**



## (B)

L.major	MN--FLGGSL----RSGGGHREKGFSAK-----NKVQQYHGDGSESIVLGRPFR
T.brucei	-----MGS----DSKSSKSKDNFSSK-----GKIEKYNKDGTTETILIMGRPFR
T.cruzi	MSTFRMAQAR---ITGRKSNKDKSFSSK-----GKIEKYNQDGTETIIVMGRPFR
G.intestinalis	-----MSHTLR
Ch.reinhardtii	MKK-----LFGGNKKKEKGTLVGKT-VM
H.sapiens	MKKFFDSRREQ-----GSGSLG-----SGSSGGGGSTSGLSGYIGRVFG-
C.owczarzaki	MKKFLSSAANAVNTISSSAGRGGGVFSSAAGGASANSSSLVGGAAADDSGAQAFIGRTISM
C.elegans	-----MAFARA---LGVIQSVGRGAGISEQSELYDRGQTFS
S.cerevisiae	-----MNQPQIG-TYNVGTQLT
D.discoideum	-----MGNNQGKTLK
P.falciparum	-----MLKFRSICTTLIGGKVYN
P.tetraurelia	-----MNF-----FMNLFQGSSGPQESYE
L.major	VLQKISDVPP-PYMTVRQRSTKYEGGSSVYLVVENTSHLPE-FPRHLALKRSFFG-VDQ
T.brucei	IVQRLSEKVLNPEAKSSKSSKKAKDVYSYIVVVENIEALPM-FARYVVVKRTFFK-IDE
T.cruzi	VLQNIASENHI-SSHKERDHFKAHLRSFVFLVENTENLSE-FPRHFALKRSFFN-VNE
G.intestinalis	FGD-----KDIVVQGEIAKGGSARIYSASCA-----GEQYVLKHLSTDPKI
Ch.reinhardtii	VGP-----FSVRVEAIVGEGGFATYRCIDGK----TGLAYALKHMLA-PDT
H.sapiens	IGR-----QQVTVEVLAEGGFAIVFLVRS-----NGMKCALCRMVFNNEHD
C.owczarzaki	TGG-----SRVMVESLIAEGGFAYVFRARDLG-----NGQAVALKRVSV-SED
C.elegans	ING-----NNYRVEKVIAGGFGTVFLATNTK-----GKQVAVKIMLSHDA
S.cerevisiae	VGS-----HQVEIIKYLTSGGFAQVYSALINPPDPHSNSVACLRVIVDPKPS
D.discoideum	IGS-----YHLNFVKQIAEGGFSYVFLVKDSN----TSKHAYLKRILIRDEDE
P.falciparum	ING-----KTIREKLISEGAYSFVYMAKDLN----TNRAYTIKKTICQDKEK
P.tetraurelia	ING-----QEYHELNLIAEGGYGFIWRAIETK----TRKFCVIKKIICQSKEA
	. . : *
L.major	VMEAHKEVEIVSRIE-DKNIVRVYHSEISRS-----
T.brucei	VLLAHREPEILNRVR-DKGILHVFHSEVTRN-----
T.cruzi	VLLAHREPEVLARVS-DKNIVRVFHTVETRS-----
G.intestinalis	YDDAITEIYFHLQFNNEKCVRLHGLLLDDTYIPANFARFDEYLTINDMHSSQSHLASSI
Ch.reinhardtii	VDEVKAEAKVMARLKGHPNIRLHAVAFAFGP-----
H.sapiens	LQVCKREIQIMRDLSGHKNIVGYIDSSINNV-----
C.owczarzaki	VAKIKSEIDFMQILSGHKNIVALKDYTLSQV-----
C.elegans	TKDIDNEIDMMKKLQ-HENIIQLFDASAESR-----
S.cerevisiae	LNTLRAEVDAMRLKNNRVVYSYIDSHAACA-----
D.discoideum	LKGVKHEISIMKRLTKHKNIVKILDYHKVS-----
P.falciparum	IDMAHKEINILKSLPPHKNIVQYYGSTIIS-----
P.tetraurelia	IEQAQLELDLHRKLQ-HPNIVKCYNGVIKFN-----
	* . . ::
L.major	-----EGRLGVSIAAMEYCSNNLYHRIR
T.brucei	-----DGRLGTVAMEYCPGSLERLQ
T.cruzi	-----DGKLGTVAMEYCSNNLYRQLR
G.intestinalis	KSSQFANQEPGLNLSSALPVPVGVSHGGHDSNRRNTTKLPCTNIRLLLEYCSGGNCVKLL
Ch.reinhardtii	-----TGAETDGFMLLDFCPLTILLE-VM
H.sapiens	-----S-SGDVWEVLIIMDFCRGGQVNVLM
C.owczarzaki	-----R-AG-LYQGFMAELCTGGQVIDLM
C.elegans	-----SSNRSVKEYKISMKEYCKFSA-DVL
S.cerevisiae	-----MLHNGSYEVFVLMYECERGLIDFM
D.discoideum	-----DKNNTMFIIMEYCSGGHVEIM
P.falciparum	-----ENNYKIVIMLMEYCEGRNLLNIF
P.tetraurelia	-----K-KLNQTIAYMVLECEGGTLIDL
	: :: *
L.major	SSAGTGIVTRLTEAEICHVMFAVTSAVGYLHAQQPPITHRDIRPENILINNKA-----
T.brucei	QE-----TRMLESEVVQVLLALASVLGYLHSRQPPVAHRNINPNSVLIHSESA-----
T.cruzi	PD-----TRVPEGEIVQVLLGITSAGVYLHSRQPPIAHRDICPENILIHSGST-----
G.intestinalis	IKCS-EHRVFLPEPLIWRFAYNITLAILCLHSK--GCSHRDLKLENILILEPYTLDDILE
Ch.reinhardtii	Q-RN--NFALDDFLVYEVFQDVVWAVAHMHKCNPPLAHRDLKAENVLKNG-----
H.sapiens	NQRL---QTGFTENEVLQIFCDTCEAVARLHQCKTPIIHRDLKVENILLHD-----
C.owczarzaki	NSH---RNGIPEAEVLRILSDVVEAVAFMHRSRPPIIHRDLKVENVLIAS-----
C.elegans	LKYK-----EVSIDFVVRIIYFTTRALVYLHSV--GAHRDIKAENLLING-----
S.cerevisiae	NTRL---QNRLHEFEILQIMSQVTQGVAAAMHALQPPLIHRDIKIENVLISA-----
D.discoideum	QKRL-SSGSKFTDQEIILKIFQDICESVAYMHSQQPLIHRDLKVENVLLDEE-----
P.falciparum	QKNK---DKIKEIHIVNILKDIITGLCFLHNQEIPIIHRDIKLENILCDK-----
P.tetraurelia	KRYN---EKRLSEQQVLLVLKQLVQAICYLHTQQPPITHRDLKVENVLLH-----
	. : . : *
L.major	-----GPTAYKLTFNGNATTEAYQCE-----TREEANMAIADIQMHTNPAFRAPEM
T.brucei	-----GASAYRLYNFRSAMTEAYHCE-----NHEEVVAAMEDFARYTTAGYRAPEM
T.cruzi	-----GAASYRLCNFSSATTEAYHCA-----NREEVAMAVEDIERNTSPGYRAPEM

G.intestinalis	NNGCTTKFPEFLRLCDFGSCTMESYDSERLRDMMENEQGMKELMRSLELKTTPCFRAPEQ
Ch.reinhardtii	-----EGRWVICDFGSATSRAQVYD-----SSADIAAEEDVIRRTTTPAYRAPEM
H.sapiens	-----RGHYVLCDFGSATNKFNPNQ-----T-EGVNAVEDEIKKYTTLSYRAPEM
C.owczarzaki	-----PGHYKLCDFGSATDKAVVPK-----S-PHIPDIEEDIQKHTTIQYRAPEM
C.elegans	-----NGKLLKLCDFGSATTKSIEMA-----PLSNSESLAVQEEFMKYTTPITRSPEV
S.cerevisiae	-----NNKYKLCDFGSVCGIIRPPR-----NSQELSYVQQDILKNNTAQYRSPEM
D.discoideum	-----SGIYKLCDFGSATEEITRMK-----NKTEMQNAEDDISRHTTLQYRAPEI
P.falciiparum	-----NGVYKICDFCSHSISNSFFP-----NDLSKSELSNLKEEIERDTTLIYRPPPEL
P.tetraurelia	-----NKVFKICDFGSASTENIDLKY---CQSNKHQISQYEENFAKQTTEIYRPPPEM
	: : * . . : * . * **
L.major	ADPWSKKRICEKTDWMSGVLLYYMMYMSLPFDPST-TIVKDWALTFPSE-----
T.brucei	LDPSSHKRIDEQVDMWALGVLLYYTMYQRLPFTDACWGLARKPKLRYPVE-----
T.cruzi	ADPGSGKRIDERVDLWSIGVLLYYMMYLRLPFGETNMGLTGRLLKLRFPAG-----
G.intestinalis	IKIDPEFPLGTAVDVFAFGILLYRLMYRRLFPDNDLRYRYRAEFSFQD-----
Ch.reinhardtii	WDLYTRQRIDTAVDVWALGVLLYVLAFGKLPFPGDSKLAILYGNESGGSNGAGAPPDAP
H.sapiens	VNLYSGKIITTKADIWALGCLLYKLCYFTLPFGES-QVAICDGNFTIPD-----NS
C.owczarzaki	IDLYSGKCIDTKADIWALGCMFLRLCYFVTPFEES-SVRIISGTYTTPA-----TP
C.elegans	CDVYSNWPIGKQDNWAMGCLLIYFVAFGEHPFD-GSALAIINGKYKKPP-----
S.cerevisiae	IDTFRGLPIDEKSDIWALGIFLYKLCYYTTPFEKGGDLAILSGKFEFFPLY-----
D.discoideum	VDFYRSPVINEKIDIWALGCLLYKLLFYVTPFEDSGSLGILNSNYTIPPN-----
P.falciiparum	IDLYSNFEISWKIDIWMVGCILYLLLFKIHPPFQNNFLSIINATFTIPYC-----
P.tetraurelia	TDLYLKYEINEKVDIWMGLCILYTMCFYNPPFQESSKLAIVEASYVIPKE-----
	. : * : . * : : : * **
L.major	---ATASYTPSLRVI-TEHLLDPDPCSRW-----DVFALINYMRFDEDCSRHLGTFC-
T.brucei	---AEVWYTGSLRIV-LEHLLPEPDPEKRW-----DAFALINMRFDNDLCRHIGFFC-
T.cruzi	---TEMWYTGSLRVV-LAHLLPEPDPKRW-----DIFALTNFRLDDEIGKHIGTFF-
G.intestinalis	---KRVKYS DALKNI-ISLCLVQHPKDRV-----SIFEIHKLVREVNRLGFIETNI-
Ch.reinhardtii	TVPHPHYPAPMPPPQYPYPQHPHPQAAQAHGAVARVPVQPPQPDLLGDACPEPPVT-
H.sapiens	R-----YSQDMHCL-IRYMLEPDPDKRP-----DIYQVSYSFSKLLKKECPIPNVQ-
C.owczarzaki	V-----YSDPIHRL-IKYMLDVDPDYRP-----NMTEVAHAVFTVRGLQSPIPLVA-
C.elegans	---VQQNQLSAFADL-IAKCLTPNPDERI-----TAAKIEEYMKLAMQKPKLAAKTD
S.cerevisiae	---PN--YSEQLKGL-IRDLVQDPRHRP-----NVYQLLKRSIMQNVPC----P--
D.discoideum	---HT--HSNDLISL-IKIMLNPDPIINRP-----NIFEITNQLNLLRNQQPLFPSH--
P.falciiparum	---TK--YSKRIISI-LLMTLNKNPQKRI-----DTSTLLFILENYADLKKWVFHI--
P.tetraurelia	---NK--YSNKMNL-IGIMLQPNPKDRA-----SIFQIEDILKNYDTLTQIVINN--
	: . *
L.major	FTKTEWPEGWEEQDVKVLGRTPPPKAPPVSYNEPGHEQLDSRGLRGQERVHGHQPP
T.brucei	FSQTERPEGWEPQDVKVIGRPVPAKVLPERLRNTGQSDDPAAAARDNRDEAAAT----S
T.cruzi	FAQTEWPDGWEQQDVKVVNQVPRKAPPVIRVNTPKHADNTKNATDAGTKNRMQA----S
G.intestinalis	FSTTGQSN-SLARSTRVSI LHAAP EA-----AHPPATENTVNIV-----
Ch.reinhardtii	-ARLQP-E-RA-----EVARLAAHNAALE-----
H.sapiens	-NSPIP-A-KLPEPVKAS-EAAAKKT-QPKARLTDPIPTTETSI-AP----RQ-----R
C.owczarzaki	-QQRGE-----RPVQPV-QSA-----PVKLQVTPIASTASPVQSR----RE-----R
C.elegans	FTDILD-L-MNVQPVAE-QSIESQAAGFFTMQDKLFSNLTSLKNTVV--QQ-----T
S.cerevisiae	INDIQ-----VVQAP-SSH-----LNLQLHQLSATQNILSLN-S--PT-----T
D.discoideum	KSNIL-----LSYNE-----NNNINNNNNNNINNN-N--NN-----N
P.falciiparum	PADIK-----KRV-----NQIFEKMNELNMN-N--KH-----T
P.tetraurelia	INDQD-----SK-----NQQEDDFGEFQSG-D--QK-----T
L.major	DDRGNGRDSAVVDGAAVQEA MIVLGGDPADDDPAVAKYRHQIIMVQOEAWRRAKAAAGLR
T.brucei	VKSRTAAQESGDNQMV LKAVTALSSDTASTDPEVLAYREKLIREQEEAWQVAESAHKKR
T.cruzi	SEMERGSSMGGDVGELDAALALGLDGD SKDPEMRKYKMLIQEQEDAWRMAKSASQAS
G.intestinalis	-----LAYKPMIPKTVHGD L-----TVVL
Ch.reinhardtii	---GRVRQ-----LEGLVAAQGSALQRLAADL-----SA--
H.sapiens	PKAGQTQP-----NPGILPIQPALTPRKRATV-----QP--
C.owczarzaki	PKASA-----
C.elegans	NKMGWGME-----
S.cerevisiae	MEN--TMP-----NATF--QISM-----
D.discoideum	IVNGKNIP-----KP----LPKV-----
P.falciiparum	D---LIE-----VS-----HEKL-----
P.tetraurelia	Q---E-----QQ---Q-----
L.major	DSPPTPEQESSAAP-----A-----PT-RANN----EAQKDV
T.brucei	DKEDKDGATSGPADVDLSFGPTEKKEAVKE-NKASAIDDLFGGME-TQPS----QSQKPT
T.cruzi	AKNKS G-DTNNVNRQ-----AQETREEKQREEEPPKQIQ-----D-NPPN----DNTPNL
G.intestinalis	DKEDSGGLDLSQASMGNLSLLEDTIASV---TSASQ---TVTGIDTGQSLSHSK----YA
Ch.reinhardtii	-----T-----SER---QVSG---SAPY--PP-----
H.sapiens	-PPQAAGS-----SNQPGLLASV---PQPKP---QAPP---SQPL--PQ---TQ
C.owczarzaki	-----PATLVSF---NQ-----QAGP-----
C.elegans	-----PTN---TTPRP---GHPS---TSPKLVPKTHRSNQ
S.cerevisiae	-----ADN---TTAQ---MHPN---R-----
D.discoideum	-----VSQ---TTPTP---TPPP---P-----

P.falci parum	-----IDV---KISSQ---KYE-----
P.tetraurelia	-----QKTV---QYD-----
L.major	FDDL FASASPTVFPA-SQAS--PAMGPSAPPATHQPPE-----KDV F-----
T.brucei	TDDL FGGASTATAS-----NQIPPQGSQV GANDSWKDDLFAAPP-----Q
T.cruzi	FDDL FAPS PAAQASS-GS-----LQAPSQPTVSRNPVTGGQNWPDSLFSGVPPSQPLSQQ
G.intestinalis	FTRSLAGSVGVNINLGLSASGV PACGDRGH-----TGARDRLTVSAIS----GLA
Ch.reinhardtii	-----
H.sapiens	AKQPQAPPTPQQTPS-TQAQGLPAQAQATP-----QH-QQQLFLKQQQ---QQQQ
C.owczarzaki	-----ASQ-QPF-GSAQ-----
C.elegans	QQEKNAPPRPPTTAS-PRAA-----
S.cerevisiae	-----KPSQI-----A-----
D.discoideum	-----APSQSPS---PS-----
P.falci parum	-----TNL---PK-----
P.tetraurelia	-----DSL---I-----
L.major	-----SASDLFSTSAPQLDPSYPVMNTYQAP-----
T.brucei	QPQM QPQPAFMAADSSWSTGAPT SVPMNPTGGMMGQQQAPMGW-----
T.cruzi	QPPLHAQNSYDMMDSGWGTGAPTMAPVG---GAYDV GAMYAG-----
G.intestinalis	DPR---AAYPGRFMQSPQCDVGLSESILMALGTP EAPHLVEELRNSIS-----
Ch.reinhardtii	---I-GYEDLT-----
H.sapiens	PPPA---QQQPAGTFYQQQQAQTQQFQAV-H---PATQQPA-----
C.owczarzaki	-PFA-GTQEQA AFYQQQQQQQMF AAQ-QQMMHMSLQHHA-----
C.elegans	-----PVATFKVSWPTDSSQNDSA-NDWGD F SQAPP AVNV-----
S.cerevisiae	-----YDASF SNSAKGSQPLFDKS-QNMYHALDPPLVEPLASSVSNNDNELKANSA
D.discoideum	-----PSPTVNNIENN SNGLEHS-NSNGNISQPSPTPPKRR-----ATPG
P.falci parum	-----SNNNFL-----KFFTTS-NLSGDI FKKKIEKEE-----
P.tetraurelia	-----
L.major	-----QQMAMGGWGS GAPVMGSGGYPMQQ-
T.brucei	-----GNGVTAPGFQ-----PMQTPQMQQQVPGGWNGATASG--FQPMQT-
T.cruzi	-----TQMATAPTFS-----MPPHPPL-QQITPSWGISAPPG-----
G.intestinalis	-----DGLTN---AC-----
Ch.reinhardtii	-----
H.sapiens	-----IAQFPVVSQ-----GGSQQQLMQNF
C.owczarzaki	-----PAQHHP LPH-----SHHQQQQQQ--
C.elegans	-----SKKDVDFPFSESASHNLSA---KHSE-----
S.cerevisiae	TKLKQAI VSEAHTFRQNN SIDFPLQNIIPQY---ED-----SSSSSDES
D.discoideum	T-----TPSLQPV SFPPNNSNNSFDDPFRDSPRTNLSN---NPFNVNSNDNS
P.falci parum	T-----KKEIKQVGN-----KQVSETEKEPKDKLEN---NPDGNKLQNDK
P.tetraurelia	-----
L.major	--PQNQQ-----QAPWGYGVSNANNYAPPLPPQ-----QQQQQLQP---F
T.brucei	--PQMQQ-----QVPGGWNGATA---PSFQPM-----QTPQMQQQVPGGW
T.cruzi	-----PQSTPM-----QQQQQQQP---F
G.intestinalis	-----ITDQQALK-----LEQVE---
Ch.reinhardtii	-----
H.sapiens	YQQQQQQQ-----QQQ-----QQQLATALHQQ---QLMTQQAALQQKP-TM
C.owczarzaki	-QQQQQQQ-----QQQ-----QQQYHEHASEH-----ASHA-TF
C.elegans	-----DLLSLQ-----
S.cerevisiae	YSGDVDE---LKKTRSLGSYS-----TRGNIKKNQSV-----
D.discoideum	NSS-----NNNNNNNNNNNNNNNNNNNNNITNEEILNII TQLTNNDQTN-TF
P.falci parum	DSGKLQNDNKLQNDKDSGKLQNDKVS GKLQNDRDGNKLQNNHII DYTDNKMNDNKKN-KS
P.tetraurelia	-----
L.major	TNF-----L-----PP-----LSGQQPGQQQS
T.brucei	GNG-----ATAPGF--QPMQT-PQ-M-----QQQVPGGAPGFMGG
T.cruzi	NHQ-----TLPTT--SPVKVSPPAIPEDSLAFKAYSLLPE-TIK-----
G.intestinalis	-----
Ch.reinhardtii	-----
H.sapiens	AAGQQPQPQFAAAPQPA PAQEP AIQAPV-R--QQP-----KVQTT P PPAVQG
C.owczarzaki	GSDQQSA--Y-WAAHRGMTQSAPIVAPV-A--TLV-----QFDSAPPAQPS
C.elegans	-----NWPGTAEHKPEVRSQ-SSRDL--FDFDDL M-----LRHTT P SAE-S
S.cerevisiae	---KES--LTSSSLPGTSFTPTSTK-----VNLKHENS PFK-S
D.discoideum	DSGLLLK--LKS-MKPGKGTMHII VKRP-LKEPLVCFK--SLLL VHALLSEGNNIQFK-S
P.falci parum	DE-ISEN--IT--SSGDKQNSVPDQSE-NQSKSHSYNLIKI QSKDTLI IDHVDRDFP-S
P.tetraurelia	-----
L.major	QQQETSQP-----
T.brucei	NDLFQR PQ-----
T.cruzi	SLPTAVPA-----

G.intestinalis	ERVQTIT-----LY-----ESLSRLQVN-----IGTLRL
Ch.reinhardtii	-----
H.sapiens	QKVGSLTPP-----SSPKTQRAGHRRILS-DV-----
C.owczarzaki	ATFDSSDPFSDMPAFAPTASDFVPVSSHVPSSAQHLTA-IP-----
C.elegans	SQAV-LQ-PIR-QVENKNLTKVDVSK-NGIGSSSSASLDDMVSDMMKMTKK-----
S.cerevisiae	TFVNTIDNSKD-DLNKPSYEDLDVSK-QSLKNSIQQRMIDKLNSEESFNARKMSKVKLH
D.discoideum	DVHD---SKD-LFNN-LYLGW--SK-Q-----KD---RYL
P.falciparum	DKTN---TDF-SFQD-TLCEF--EK-K-----
P.tetraurelia	-----
L.major	-----R-----
T.brucei	-----Q-----
T.cruzi	-----
G.intestinalis	PLLPPLL-ETANLFKI-----
Ch.reinhardtii	-----
H.sapiens	-----TH--SAVFGVPASKSTQLLQAAAAEASLNKSKSATTTPSGSPRTSQQNVYNPSE
C.owczarzaki	-----AH--LHLFSPQTDHSPGP-R-SAPPHSFPGPPPAATTQPGA---PPQVTIAPPP
C.elegans	-----
S.cerevisiae	EKGEID-----KPIMLKSSGP-----ISKDKKTK--PTPPPKPSHL--KPKPPP
D.discoideum	QLGELLSHYSLLLYKFILFHQK-N-----YMIDGSFAFEEMKWGIPESL--DSNNHP
P.falciparum	-----CMNNS-N-----YSNDKKVA-----GDI--KKDDCS
P.tetraurelia	-----
L.major	---TARATTKRDPFADLFN-----
T.brucei	---QQPQQPEKDPFASLFK-----
T.cruzi	---EPKPEAKKDPFADLFL-----
G.intestinalis	-----
Ch.reinhardtii	-----
H.sapiens	G-----STWNPFDNFSKLTAEELLNKD-F--AKLGEGKHPEKLGGSAESLIPGFQ
C.owczarzaki	S-----STFAPLNROTSSLANWDGAPARLP--AHLGH---RRIGSSPADIRFGHH
C.elegans	-----
S.cerevisiae	KPLLAGRKLSLDK-----
D.discoideum	--ISINTIKALFDIMDHLFLVQNNLSDYCINNICSNSSSGSGNVPISSLQHCVN-----
P.falciparum	--VNI-----QVVNNNDVCHND-----MKDVNECTN-----
P.tetraurelia	-----
L.major	-----
T.brucei	-----
T.cruzi	-----
G.intestinalis	-----
Ch.reinhardtii	-----
H.sapiens	--STQGDAFATTSFSAG-TAEKRKGGQ-----TVDSG--LPLLSV-----SD
C.owczarzaki	APPAAGEPFAHS-FSGESVTSSSSAGS-----TWGADFLSPDSST-----NR
C.elegans	-----
S.cerevisiae	-----
D.discoideum	--ILNSSSYSIFCFISGSIDV-----
P.falciparum	--LLNLDVFEYENETNDSIKKINDGDFIKQSGCDFIKQNGGDFIKPNDSNCIFADINFNN
P.tetraurelia	-----
L.major	-----
T.brucei	-----
T.cruzi	-----
G.intestinalis	-----
Ch.reinhardtii	-----
H.sapiens	-----PFIPLQ-----VPDAPEKLEGLK-----
C.owczarzaki	-----SISPLQ-----EEDNAK-LVQL-----
C.elegans	-----
S.cerevisiae	-----
D.discoideum	-----LS-KQFSDVDMKLISCNVQFQSMYTRLRE
P.falciparum	DINLNNDINLNNDIDFFDVIQNHDLKKKEDVLNINRNNEHDKNAITCSNEYYNLS-----
P.tetraurelia	-----
L.major	-----
T.brucei	-----
T.cruzi	-----
G.intestinalis	-----
Ch.reinhardtii	-----
H.sapiens	-----SPDTSLLLPLDLLP---MTDPFGSTSDAIVIGKVIISVSSVMHDMCACFKNDKY
C.owczarzaki	-----
C.elegans	-----
S.cerevisiae	-----
D.discoideum	QYTKLAQVPCFSDIFFPTLPNTAPTFTIVRSNSFNR-----LNSSLSDLNLNNNNNNN



P.falciparum	-----LFYDPWNNNTKDNHFFGDDDYNVNKSSSFI FNNDLFDNYELCTNNNTE
P.tetraurelia	-----
L.major	-----
T.brucei	-----
T.cruzi	-----
G.intestinalis	-----
Ch.reinhardtii	-----
H.sapiens	LVNQSLGNSPA-----TPEAKAI-----
C.owczarzaki	-----
C.elegans	-----
S.cerevisiae	-----
D.discoideum	NNSNNSNNSNNSGNSGNASLNSSFD---NINSSNPFSTPTFNPFSATTTNTSESG
P.falciparum	KSTHHVKNNCMDINNNIFSNKNYDEHVDKKEKNISYDYNVKSHPITYSSYKNEK----R--
P.tetraurelia	-----
L.major	-----
T.brucei	-----
T.cruzi	-----
G.intestinalis	-----
Ch.reinhardtii	-----
H.sapiens	-----
C.owczarzaki	-----
C.elegans	-----
S.cerevisiae	-----
D.discoideum	FGNFGSLSEPTSNPSPRYQQSNNNNNNNNNNGTPISLTPGSLSPVIGAKKPPLPPNIHHL
P.falciparum	---CNLNDIKQNNSSKYIMSDKND-----KFSELL-----DEF
P.tetraurelia	-----
L.major	-----
T.brucei	-----
T.cruzi	-----
G.intestinalis	-----
Ch.reinhardtii	-----
H.sapiens	-----
C.owczarzaki	-----
C.elegans	-----
S.cerevisiae	-----
D.discoideum	QQQQHPQQQQQQQQQQQQQQQQQQQQQQQQQQQQHPQQQGLRFPHSASLEDARLYTLILTP
P.falciparum	QKCT-----IK-----
P.tetraurelia	-----
L.major	-----
T.brucei	-----
T.cruzi	-----
G.intestinalis	-----
Ch.reinhardtii	-----
H.sapiens	-----
C.owczarzaki	-----
C.elegans	-----
S.cerevisiae	-----
D.discoideum	SSSPPLSPSTGPTTAAQQQQQQQQSQHTFDNFNINGHAPPVPQSTQPSFQPHVSFAPNVN
P.falciparum	-----
P.tetraurelia	-----
L.major	-----
T.brucei	-----
T.cruzi	-----
G.intestinalis	-----
Ch.reinhardtii	-----
H.sapiens	-----
C.owczarzaki	-----
C.elegans	-----
S.cerevisiae	-----
D.discoideum	INNNNNSHVSAPHSLNSSSSSISSISNPNLGGIAQKSGNSLMPPPLYKPAARGHRRSQS
P.falciparum	-----
P.tetraurelia	-----
L.major	-----
T.brucei	-----
T.cruzi	-----

G.intestinalis	-----
Ch.reinhardtii	-----
H.sapiens	-----
C.owczarzaki	-----
C.elegans	-----
S.cerevisiae	-----
D.discoideum	SNGDEVRRRNLLQQQLEQNRDFLNHNRLLNKQSRMNNPNNLFDEGDSGFGDGEEDEGLL
P.falciparum	-----
P.tetraurelia	-----

L.major	-
T.brucei	-
T.cruzi	-
G.intestinalis	-
Ch.reinhardtii	-
H.sapiens	-
C.owczarzaki	-
C.elegans	-
S.cerevisiae	-
D.discoideum	N
P.falciparum	-
P.tetraurelia	-

## (C)

L.major	MN--FLCGSLRSGGGHRREKGFSAKNKVQQYHGDGSESVLAGRPFRVLQKISDVPP-PY
T.brucei	-----MGSDSKSSKSKDKNFSSKGKIEKYNKDGTTETILIMGRPFRIVQRLSEKVLNPE
T.cruzi	MSTFRMZAQARITGRKSNKDKSFSSKGKIEKYNQDGTETIIVMGRPFRVLQNI GASENHI-
	: :*:*:*.*:*:*: *:*:*:*: *****:*.:

L.major	MVTRQRSTKYGEGSSYVVLVENTSHLPEFPRLHALKRSFFGVDQVMEAHKEVEIVSRIE
T.brucei	AKSSKKSSKKAKDVYSYIVVVENIEALPMFARYVVVKRTFFKIDEVLLAHREPEILNVR
T.cruzi	SSHKERDHKFAKHLRSFVFLVENTENLSEFPRLHALKRSFFNVNEVLLAHREPEVLARVS
	:. . * .: :*: :*** . * * * :. . :*:*:*: :*: :*: * *: *

L.major	DKNIVRVYHSEISRSEGR LGVSIAMEYCSNNLYHRIRSSAGTGIVTRLTEAEICHVMFAV
T.brucei	DKGILHVFHSEVTRNDGRLGTTVAMEYCPGSLEQRLQQE-----TRMLESEVVQVLLAL
T.cruzi	DKNIVRVFHTEVTRSDGKLGVTVAMEYCSNNLYRQLRPD-----TRVPEGEIVQVLLGI
	**:*:*:*:*:*:*:*:*:*:*:*:*:*. . * :. . :. : * *: * *: *: *: *

L.major	TSAVGYLHAQQPPITHRDIRPENILINNKHAGPTAYKL TNFGNATTEAYQCETREEANMA
T.brucei	ASVLGYLHSRQPPVAHRNINPSNVLIHSESAGASAYRLYNFRSAMTEAYHCENHEEVVAA
T.cruzi	TSAVGYLHSRQPPIAHRDIPCENILIHSGSTGAASYRLCNFSSATTEAYHCANREEVAMA
	:*:.*:*:*:*:*:*:*:*:*. . * *:*. . :* :*:*: * * . * *****:*. . *

L.major	IADIQMHTNPAFRAPEMADPWSKKRICEKTD MWSMGVLLYYMMYMSLPFDPST-TIVKDW
T.brucei	MEDFARYTTAGYRAPEMLD PSSHKRIDEQVDMWALGVLLYYTMYQRLPFTDACWGLARKP
T.cruzi	VEDIERNTSPGYRAPEMADPGSGKRIDERVDLWSIGVLLYYMYMLRLPFGETNMGLTGRL
	: * : * . . :***** * * * * * *:*:*:*:***** * * * * : :.

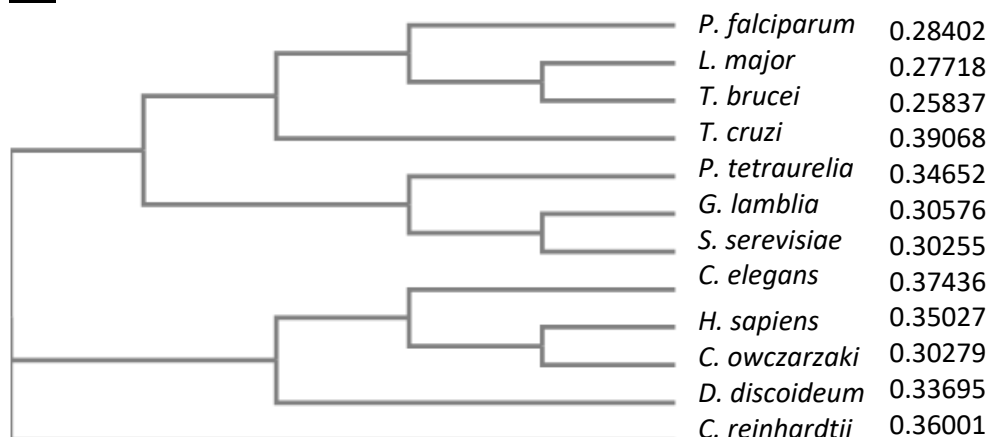
L.major	ALTFPSEATASYTPSLRVITEHLDPDPCSRWDVFALINYMRFDEDCSRHLGTFCTKTE
T.brucei	KLRYPVEAEVWYTGSLRIVLEHLLPDPEKRWDAFALINFMRFNDLCRHIGPFCFSQTE
T.cruzi	KLRFPA GTEMWYTGSLRVVLAHLLPDPKRWDIFALTNFLRFDEIGKHIGTFFFAQTE
	* : * : * * * *: * *:*. . * * * * * *:*:*:*: :*: * * *:*: *

L.major	WPEGWEEQDVKVLGRTPPPKAPPVSYNEPGHEQLDSRGLRGQERVHGHTNGHQPPDDRGN
T.brucei	RPEGWEPQDVKVIGRPVPAKVLPERLNTGQSDDPAAAARDNRDEAAAT-----SVKSRT
T.cruzi	WPDGWEQQDVKV VNRQVPRKAPPIRVNTPKHADNTKNATDAGTKNRMQA-----SSEMER
	*:*** *****:*. * * . * . : : . : .

L.major	GRDSAVVDGAAVQEAMIVLGGDPADDDPAVAKYRHQIIMVQQEAWRRAKAAAGLRDSPPT
T.brucei	AAQESGDNDQMV LKAVTALSSDTASTDPEVLAYREKLIREQEEAWQVAESAHKKRDKEDK
T.cruzi	GSSMGGDVGELDAAALALGLDGSDKDP EMRKYYKMLIQEQEDAWRMAKSASQASAKNKS
	. . . . * : . * . * * : * . : * *:*:*: *:*: *

L.major	EPQESSAAP-----A-----PTRANNEAQKDVFDLDFASASP
T.brucei	DGATSGPADVDSLFGPTEKKEAVKE-NKASAIDDLFGGMETQPSQSQKPTTDDLFGGAST
T.cruzi	G-DTTNVNNRQ---AQETREEKQREEEPPKQIQ-----DNPPNDNTFNLFDLDFAPSPA
	:. . : ****. :
L.major	TVFPASQASPAMGPSAPPATHQPPE-----KDVF-----S
T.brucei	ATAS-----NQIPPQGSQVGANDSWKDDLFAAPP-----QQPQMPPQPAFPMA
T.cruzi	AQASSGS---LQAPSQPTVSRNPVTGGQNWPDLSLFGVPPSQPLSQQQPPLHAQNSYDMM
	: . . . . : *
L.major	ASDLFSTSAPQLDPSYPVMNTYQQAP-----QQMAMGGWGSGA
T.brucei	ADSSWSTGAPTSVPMNPTGGMMGQQQAPMGWNGVTAPGFQPMQTPQMQQQVPGGWNGA
T.cruzi	MDSGWGTGAPTMAPVG---GAYDVGAMYAG-TQMATAPTFSMPHPPL-QQITPSWGISA
	. . : . * . . * . * . *
L.major	PVMGSGGYPMQQPQNQQQAPWGYGVSNANNYAPPLPPQQQQQLQP----FTNFL----
T.brucei	TASG--FQPMQTPQMQQQVPGGWNGATA---PSFQPMQTPQMQQQVPGGWNGATAAPGF
T.cruzi	PPG-----PQSTPMQQQQQQQP-----FNHQ-----
	* * * * * : :
L.major	-----PP----LSGQQPGQQQSQQQETSQPRTARATTKRDPFADLFN
T.brucei	QPMQTPQMQQQVPGGAPGFMGGNDLFRPQQQQPQQPEKDPFASLFK
T.cruzi	-----TTMDFAQGSLPTAVPA-EPKPEAKKDPFADLFL
	. . : : ****. *

**(D)**



**Figure 4.12 (A)** Schematic cartoons shows the domain structure of protein kinase in different organisms. Protein kinase domain is highlighted in red. **(B)** Multiple sequence alignments compared all top hit protein kinase that came back from NCBI BLAST when *Tb* protein kinase was the query sequences. **(C)** Multiple sequence alignments between *T. brucei*, *T. cruzi* and *L. major*. Red bracket indicates protein kinase domain. "\*" indicates identical aa and ":" indicates conserved aa. **(D)** Phylogenetic analysis of the protein kinase by CLUSTAL OMEGA

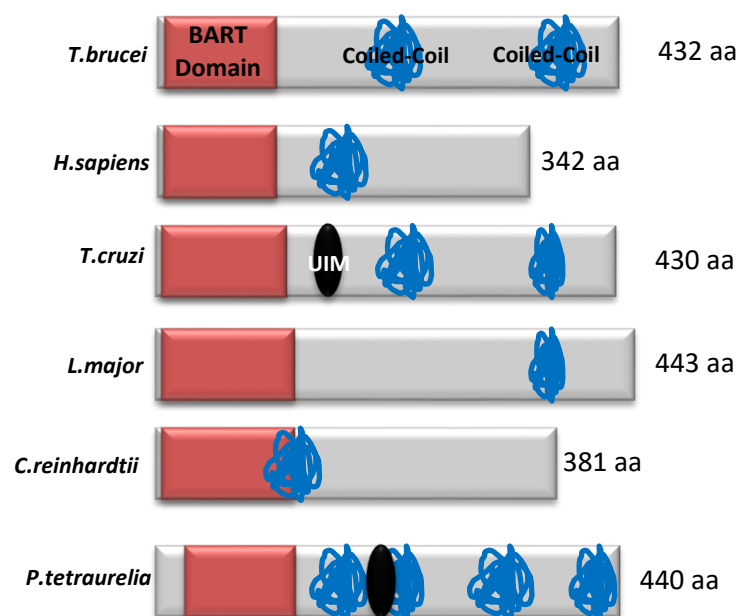
## 4.12 ARL2-binding protein (BART, Binder of Arl Two) like 1

BARTL1 (or Cilia- and flagella-associated protein 36 (CFA36), or Coiled-coil domain-containing protein 104 (CCDC104)) was identified by my *Tb*RP2 BioID-SILAC analysis using the whole cell sample. The SILAC ratio of 0.9 means the protein was only twice as enriched in *Tb*RP2::BirA\* induced cells than in the parental cell line and so was below the 1.0 cut-off value that I initially selected. However, the reason this protein was selected for further analysis was that during the time of my study Lokaj et al., (2015), identified BARTL1 as a novel Arl3-interacting protein with structural homology to the binder of Arl2 (BART). Their data showed that BARTL1 is a transition zone protein, which enhances the interaction between Arl3 and RP2 (Lokaj et al., 2015). Additionally, the BART-like domain of the BARTL1 protein can recognise and interact with a LLxLxxL motif at N-terminal amphipathic helix of Arl3 that may suggest that BARTL1 is involved in the Arl3 transport network in cilia. However, no noticeable phenotype was detected in mouse renal epithelial (IMCD3) cells depleted for BARTL1 (although it was suggested that this might have been due to the low efficacy of siRNA) (Lokaj et al., 2015). Bioinformatics analysis shows that there is only one BARTL1 gene (*Tb*927.10.5810) in the *T. brucei* genome and when I used it as my query sequence in NCBI BLAST, homologous sequences could not be identified in most of the organisms (apart from *P. tetraurelia* and *C. reinhardtii*) shown in table 4.2. Unsurprisingly, the two species closely related to *T. brucei* (i.e. *T. cruzi* and *L. major*) both contained sequences with a good e value. The top hit in the human genome is CCDC104 protein (AAH10011.1) and a reciprocal BLAST confirmed that *Tb*BARTL1 is likely a orthologue of the Human CCDC104 protein. A BARTL1-like protein was not detected in *P. falciparum*, *G. Lamblia*, *C. elegans*, or in all three non-ciliated organisms. However, the green algae *C. reinhardtii* and the protist *P. tetraurelia* encode a BARTL1 protein in their genomes. When I used the human CCDC104 protein as my query, NCBI BLAST revealed that apart from three non-flagellated organisms, *P. falciparum* and *G. lamblia* did not show any CCDC104-like protein, but the rest of the organisms all appear to have at least one BARTL1 gene. InterProScan was used to analyse the

domain structures of the BARTL1 protein (Fig 4.13-A). All proteins share a N-terminal ADP-ribosylation factor-like 2 (ARF2) binding domain (also known as BART). BART was first identified as an effector for Arl proteins, and specifically interacts with ARL2-GTP form with high affinity, but not to ARF2-GDP (Sharer and Kahn, 1999). BART is predominantly localised in the cytosol, but also appears to be associated with mitochondria, and is involved in binding to an adenine nucleotide transporter (ANT1) *in vitro* assays (Sharer et al., 2002). Coiled-coil domains can also be detected, which locate either at the C-terminus, or the central region or both locations within the protein sequence. Moreover, in *T. cruzi* and *P. tetraurelia* another motif called Ubiquitin Interacting Motif (UIM), or 'LALAL-motif' is also detected. UIMs normally form a short alpha-helix, and they can bind to ubiquitin and serve as a specific targeting signal for ubiquitination and ubiquitin metabolism (Hofmann and Falquet, 2001; Young et al., 1998).

In the next following two chapters, I undertake a functional interrogation of the proteins covered in this chapter by (chapter 5) localisation and (chapter 6) RNAi.

### (A)



[illegible]

Ch.reinhardtii	IMGAN-----
L.major	LIGEARKSSG
T.brucei	ILDETKK---
T.cruzi	IVHEATK---
H.sapiens	VINK-----
P.tetraurelia	AQQQEQS---

**Figure 4.13 (A)** Schematic cartoons shows the domain structure of BARTL1 proteins in different organisms. BART domain is highlighted in red, Ubiquitin Interacting Motif (UIM) in black and coiled-coil is in blue. **(B)** Multiple sequence alignments compared all organisms which has BARTL1 like proteins (NCBI BLAST) when *Tb*BARTL1 was the query sequence. Red bracket indicates BART domain. “\*” indicates identical aa and “:” indicates conserved aa

## **Chapter 5 Localisation of putative *Tb*RP2 interacting /near neighbour proteins**

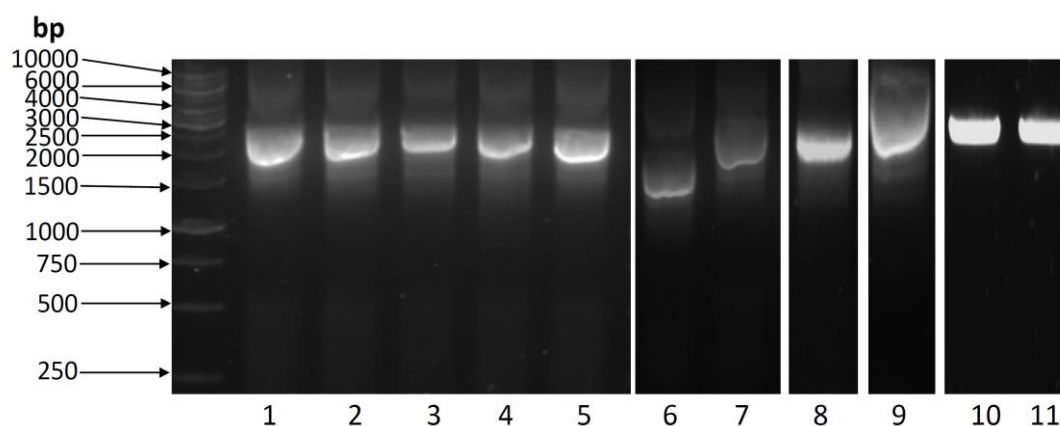
### **5.1 Introduction**

In this chapter, I describe the localisation of 11 putative *Tb*RP2-interacting/near neighbour proteins identified in the BioID experiments described in Chapter 3; with a pPOT-tagging strategy (Dean et al. 2015) employed to generate procyclic cell lines in which proteins of interest were expressed as endogenous gene fusions.

### **5.2 Generation of procyclic cell lines expressing YFP-tagged proteins**

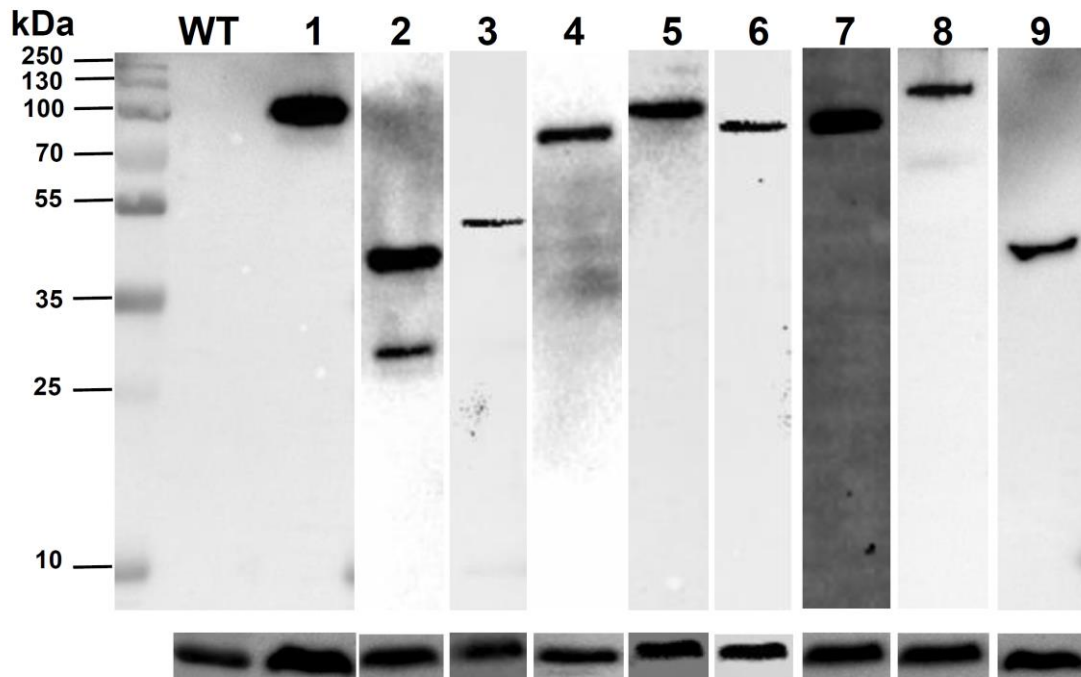
A PCR only tagging (pPOT) approach was employed to generate *T. brucei* procyclic cell lines capable of constitutively expressing YFP-tagged versions of the putative *Tb*RP2-interacting/near neighbour proteins. In brief, DNA sequence encoding the YFP open reading frame and a drug resistance cassette was amplified by PCR, using long primers that incorporated 80 nucleotides of homology to the target gene and its adjacent untranslated region (UTR) (Dean et al. 2015). Unless otherwise stated, all amplicons generated in this study were designed to allow expression of C-terminal YFP fusion proteins. As pPOT tagging amplifies the same plasmid template and incorporates ~80 nucleotides of homology to a target gene and UTR, amplicons to tag the C-terminus of a target protein are of similar size (~2500bp). Amplicons to tag the N-terminus of a target protein are smaller (~1500bp). The PCR amplicons generated for all eleven potential *Tb*RP2 interacting/nearby partners are shown in Fig 5.1. The PCR amplicons were subsequently purified and used to transfect procyclic *T. brucei*, with the gene specific ORF and UTR sequences enabling the amplicon to recombine into the *T. brucei* genome such that YFP coding sequence was fused with the gene of interest. Recombinant cell lines were obtained by using appropriate drug selection.





**Fig 5.1 PCR products of all selected candidate proteins**, which resolved on a 0.8% agarose gel and visualised with Gel red. Lane 1-5 and 7-11 were C-terminal tagged with ~ 2000bp, and lane 6 was N-terminal tagged with predicted size ~ 1500bp.

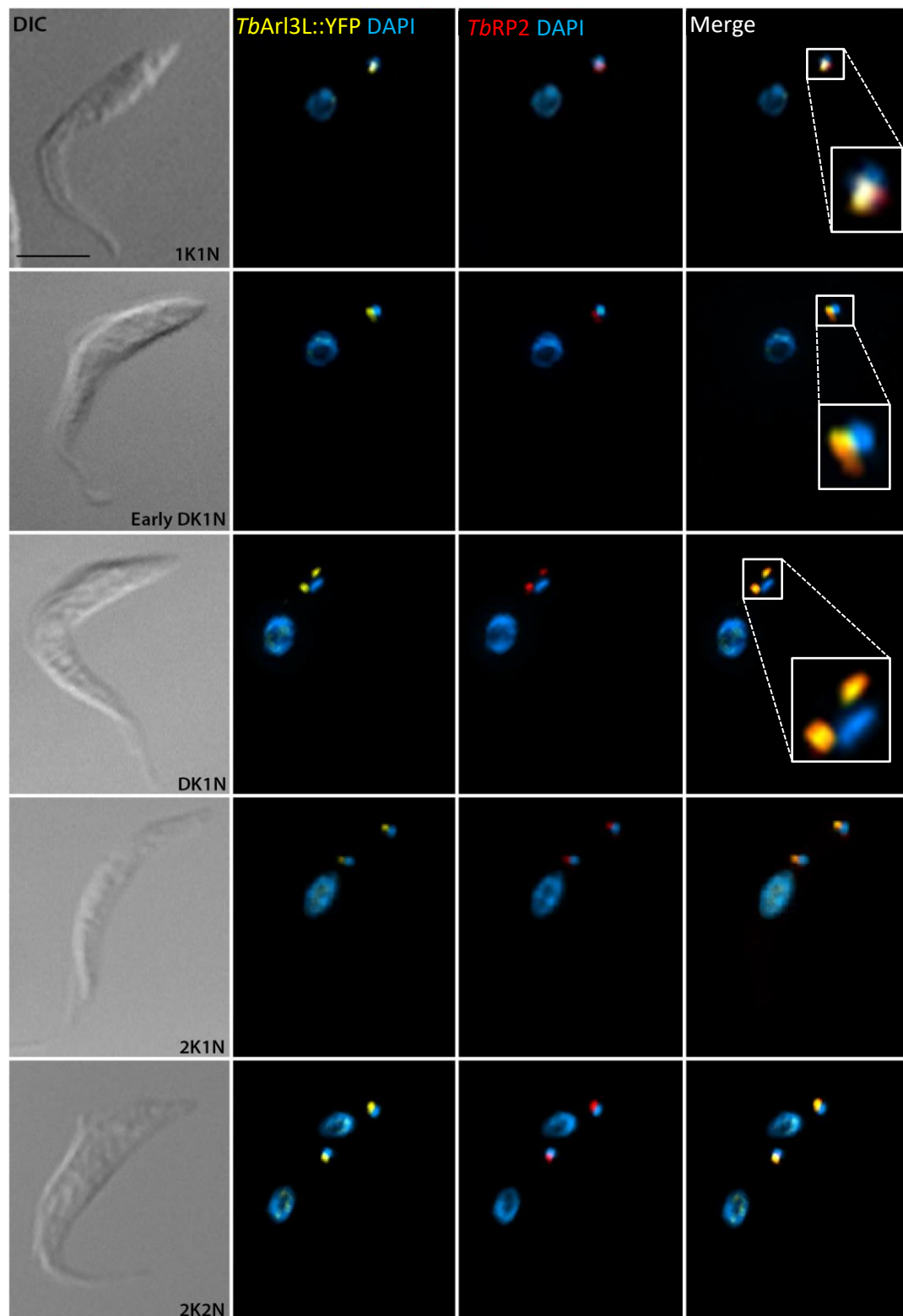
Immunoblotting was used to confirm the expression of YFP-tagged proteins in the cell lines generated. Whole-cell lysates were resolved by SDS-PAGE, transferred to PVDF membranes and probed with antisera against GFP (Fig 5.2); the anti-GFP antisera cross-reacts with YFP. The immunoblotting experiments revealed YFP-fusion proteins corresponding to the expected molecular weights were expressed in eight of the cell lines generated. However, in the cell line expressing *Tbp25-α::YFP* (Figure 5.2, lane 2) two bands were detected; the higher MW band corresponds to the expected size of the *Tbp25-α::YFP* fusion protein (~43kDa), and the lower MW band of a cleavage product of *Tbp25-α::YFP* (~25kDa). A protein of this MW cannot be detected in any other cell line generated in this study. In the immunoblotting experiments, the SmOxp9 parental cell line acted as the negative control, and the anti-β-tubulin antibody (KMX-1) used to ensure equivalence of protein loading. The immunoblotting experiments revealed variation in protein expression between the cell lines, even though the KMX-1 control indicated protein (and hence cell number) loading was approximately equivalent. For example, compare expression of *TbHSP70::YFP* (Figure 5.2, Lane 1) with *TbARL3::YFP* (Figure 5.2, Lane 3). It was not possible to visualise expression of *TbBBP590::YFP* or *TbMARF::YFP* by immunoblot analysis, due to the high molecular weight of these fusion proteins; ~600kDa and 400kDa respectively.



**Fig 5.2 Confirmation of the expression of YFP-tagged proteins by immunoblotting.** Protein samples were resolved on 10 % SDS-PAGE gel, transferred on PVDF blotting membrane, probed with anti-GFP antibody and visualised with ChemiDoc™ imaging system. Expected molecular size list as follow; Lane 1:*TbHSP70::YFP* (**98kDa**); Lane 2:*TbP25-a::YFP* (**43kDa**); Lane 3:*TbArl3 Like::YFP* (**47kDa**); Lane 4:*TbBart Like::YFP* (**75kDa**); Lane 5:*TbTCP-1-e::YFP* (**86kDa**); Lane 6:*TbVHS domain containing protein::YFP* (**76kDa**); Lane 7:*TbCytoskeleton associated protein::YFP* (**78kDa**); Lane 8:*TbProteinKinase::YFP* (**107kDa**) and Lane 9:*TbKMP-11::YFP* (**38kDa**). The expression of two large candidate proteins (*TbBBP590* and *TbMARF*) are not shown here.

### 5.3 ADP-ribosylation factor-like protein 3 Like (ARL3L)

The expression and intracellular localisation of *Tb*ARL3L::YFP was determined in whole cells at various stages of the cell division cycle (Figure 5.3). At beginning of the cell cycle (1K1N), a discrete focus of *Tb*ARL3L::YFP was detected close to the kinetoplast, as *T. brucei* progressed thorough the cell cycle (divK1N), the signal resolved as two discrete foci that segregated with dividing kinetoplasts; The *Tb*ARL3L::YFP is a detergent-extractable protein, as the YFP signal lost in cytoskeletons. To further interrogate the *Tb*ARL3L::YFP localisation, cells were co-labelled with a polyclonal antibody against *Tb*RP2, revealing the *Tb*ARL3L::YFP signal was coincident with *Tb*RP2 at the mature basal body. In addition to the mature basal body localisation, *Tb*ARL3L::YFP also appeared to localise to the nucleus (Figure 5.3). The *Tb*ARL3L::YFP signal within the nucleus was much weaker, and was not uniformly distributed but presented with a distinct punctate pattern. However, nuclear foci followed no interpretable cell cycle-specific pattern, and it remains unclear what these foci represent. For instance, there is no evidence that *Tb*ARL3L::YFP locates to the mitotic spindle or the spindle pole during mitosis.



**Figure 5.3 Immunofluorescence images showing the localisation of *TbArl3L::YFP* in procyclic form *T. brucei*.** Whole cells were co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Cells were visualised by DeltaVision Deconvolution microscopy. Scale bar = 5µm

## 5.4 Kinetoplastid membrane protein 11 (KMP-11)

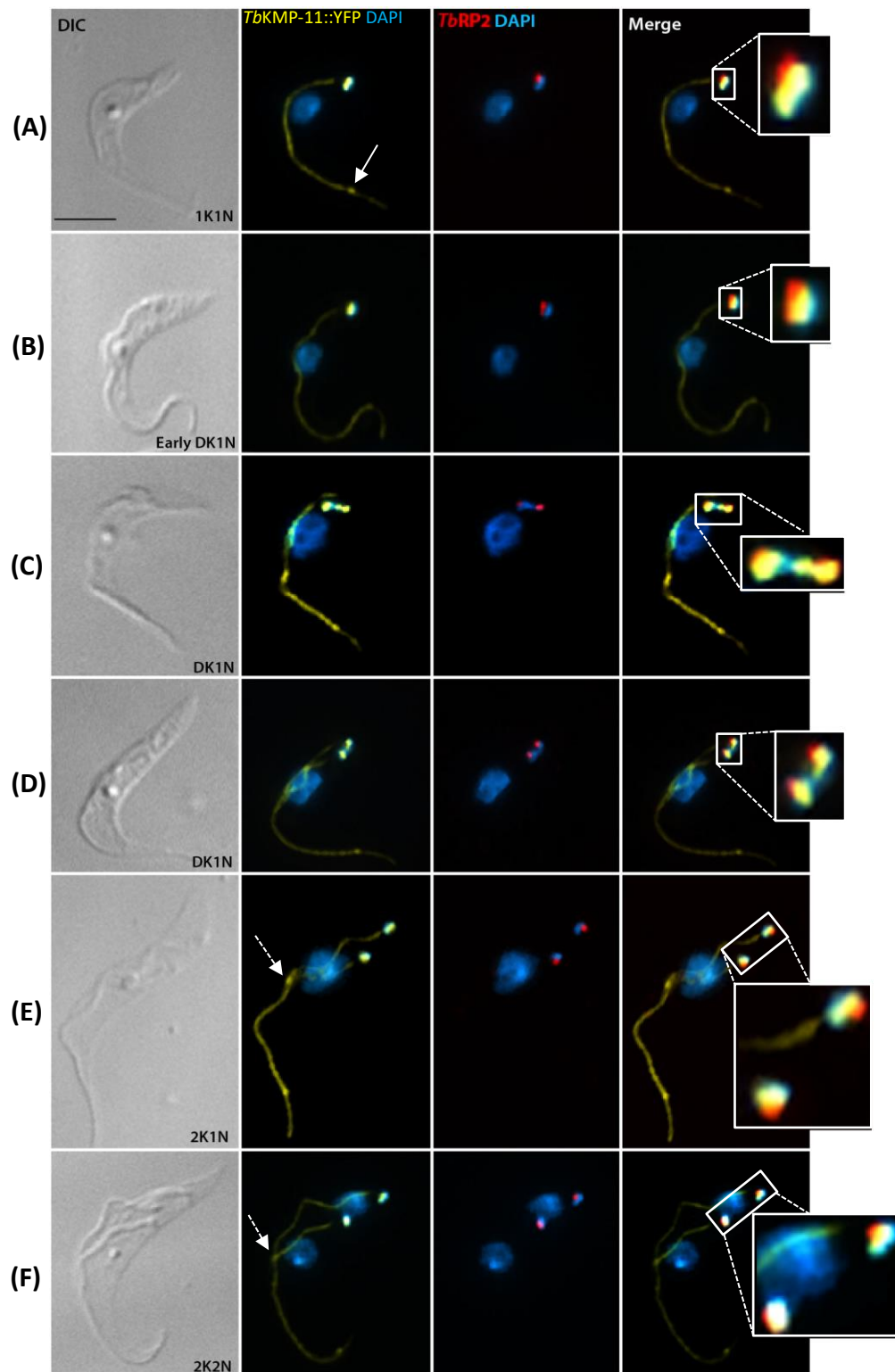
The localisation of *TbKMP-11::YFP* was investigated in detergent-extracted cells using fluorescence microscopy, revealing *TbKMP-11::YFP* at several distinct localisations within procyclic cells (Figure 5.4-(A-F)). At the beginning of the cell division cycle (1K1N), two foci of *TbKMP-11::YFP* were observed close to the kinetoplast. Co-labelling *TbKMP-11::YFP* expressing cells with the anti-*TbRP2* antibody revealed the *TbKMP-11::YFP* foci were positioned between the DAPI stained kinetoplast and *TbRP2*, or partially overlapped with the *TbRP2* signal (Figure 5.4-(A)). Later in the cell cycle (i.e. two *TbRP2* foci and two mature basal bodies) four *TbKMP-11* foci were observed; a labelling pattern that may indicate *TbKMP-11::YFP* may associate with both basal and probasal body (Figure 5.4-(B-F)). In addition to this putative basal/probasal body signal, *TbKMP-11::YFP* also localised to both new and old flagella throughout the cell cycle. However, *TbKMP-11::YFP* signal was not evenly distributed along the old flagellum, being, (i) stronger in the region where the old flagellum is attached to the cell body, (ii) present as a distinct focus at the anterior pole of the cell body, (iii) fainter in the region of the flagellum that extends beyond the cell body (Figure 5.4-(A); white arrow). In contrast, the *TbKMP-11::YFP* signal was uniform in the new flagellum, presumably, as it does not extend beyond the anterior pole of the cell. The white arrow with broken line indicated that *TbKMP-11::YFP* could be also possible has flagellar connector localisation (Figure 5.4-(E-F); white arrow).

### 5.4.1 Further interrogation of *TbKMP-11::YFP* flagellum localisation

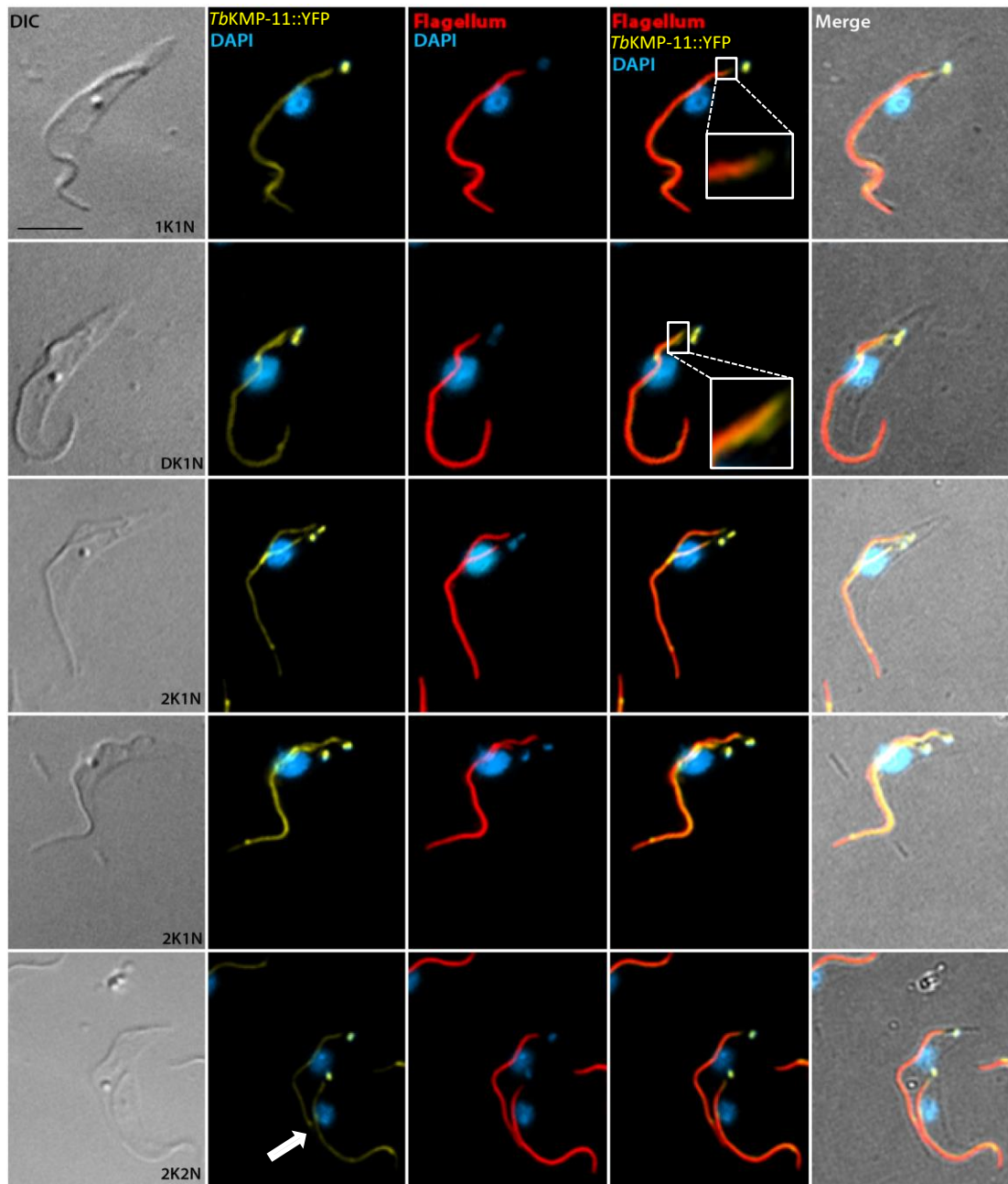
To provide further insight into the intraflagellar localisation of *TbKMP-11::YFP*, cells were co-labelled with the monoclonal antibody L8C4, which recognises the PFR protein PFR2 (Figure 5.5). Visualisation of *TbKMP-11::YFP* and PFR2 at various stages of the cell division cycle indicated the two proteins overlapped in both new and old flagella. However, it was observed that the *TbKMP-11::YFP* signal also extended to a more proximal position within the flagellum (i.e. closer to the kinetoplast) than the

PFR (Figure 5.5; Inset square). This result is intriguing since the PFR is only assembled onto the axoneme after the flagellum has exited flagellar pocket, so the proximal extension of *TbKMP-11::YFP* towards the kinetoplast may suggest *TbKMP-11* is associated with the axoneme at this proximal location. Furthermore, as the *TbKMP-11::YFP* signal within the flagellum becomes fainter as the flagellum extends beyond the end of the cell body while the PFR signal remains strong, it also suggests *KMP-11* is not integral to the PFR structure.

During the course of undertaking these various immunolabelling experiments a clear focus of *TbKMP-11::YFP* was also observed at the distal tip of elongating new flagellum. To investigate whether this foci of *TbKMP-11::YFP* could be associated with the flagellar connector (which forms at the end of the new flagellum to guide its outgrowth), cells were co-labelled with the monoclonal antibody AB1 (a defined antibody marker for the flagellar connector; (Briggs et al. 2004)). In cells with two flagella, the *TbKMP-11::YFP* signal associated with the tip of the new flagellum was co-localised with AB1, indicating flagellar connector association (Figure 5.6 Panels A-C). As the cell progresses through the cell division cycle (2K1N) and the new flagellum extends further, AB1 and a weak *TbKMP-11::YFP* signal is still detectable at the tip of the new flagellum. However, when cells have entered the later stages of the cell division cycle (2K2N), neither AB1 nor the *TbKMP-11::YFP* signal could be detected at the tip of the new flagellum.

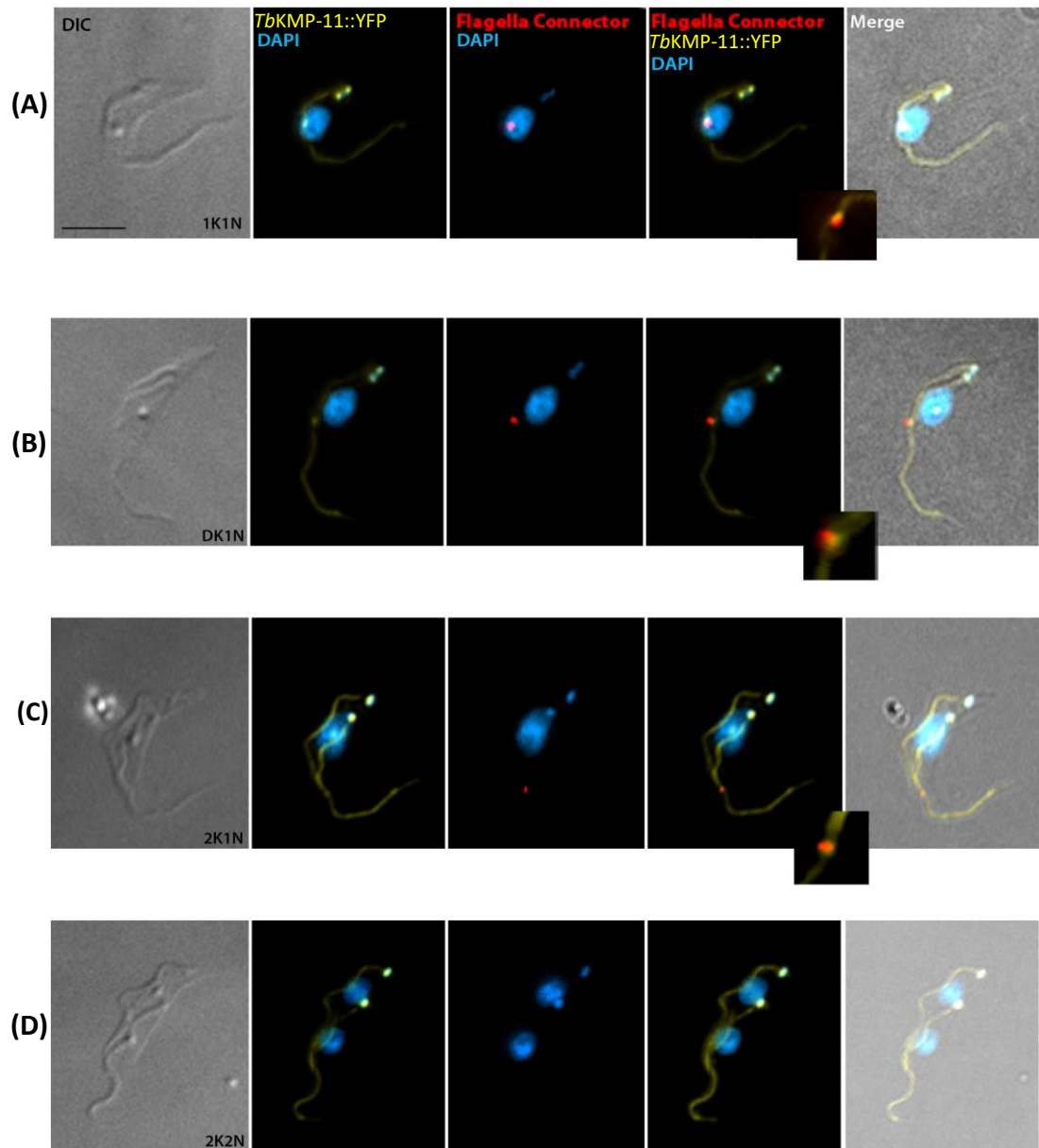


**Figure 5.4** Fluorescence images showing localisation of *TbKMP-11::YFP* in cells at different stages of the cell cycle. *T. brucei* procyclic cells expressing *TbKMP-11::YFP* (yellow) were detergent extracted and co-stained with an antibody specific for *TbRP2* (red) and DAPI (blue). White arrows (solid line) indicate intense signal at the end of cell, however beyond this point where the flagellum is no longer attached to the cell body, the *TbKMP-11::YFP* signal is noticeably weaker. White arrows (broken line) indicate *TbKMP-11* signal in the vicinity of the flagella connector. Scale bar = 5 µm.



**Fig. 5.5 Immunofluorescence images of cells expressing *TbKMP-11::YFP* co-labelled with the anti-PFR antibody L8C4.** *T. brucei* procyclic cells at various stages of the cell division cycle were visualised, revealing that *TbKMP::YFP* signal partially overlaps with the PFR and extends beyond the posterior pole of the cell (Inset). White arrow indicates the tip of the new flagellum. DAPI (blue) and L8C4 (red). Scale bar = 5  $\mu$ m.

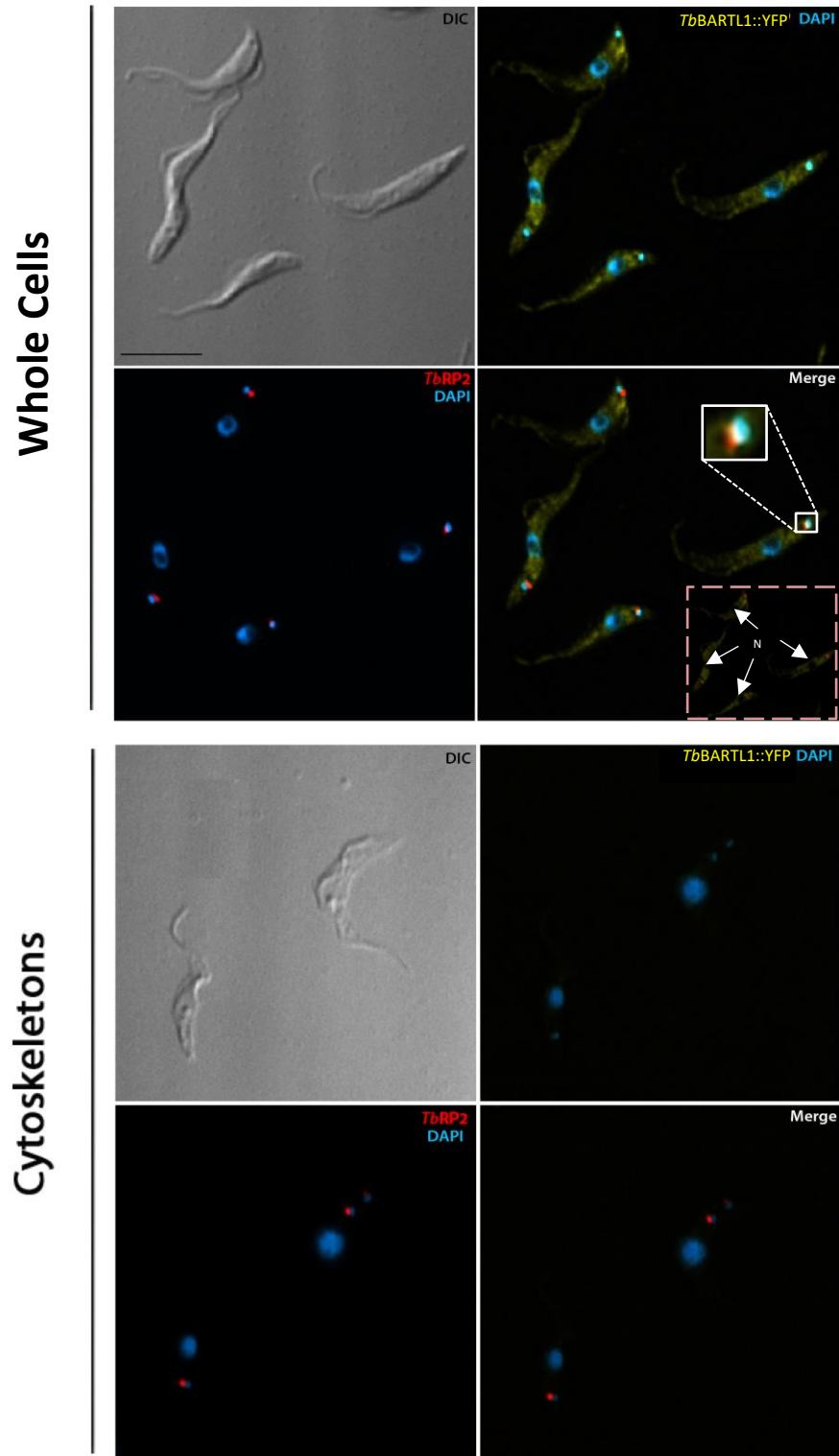




**Fig. 5.6** Immunofluorescence images of *T. brucei* procyclic cells showing localisation of *TbKMP-11::YFP* in relation to the flagellar connector antibody AB1. In cells that are elongating a new flagellum, *TbKMP-11::YFP* also co-localises with the FC, as detected by the antibody AB1. AB1 and *TbKMP::YFP* labelling is not observed at the tip of the new flagellum in 2K2N cells DAPI (blue) and AB1 antibody (red). Scale bar = 5 μm.

## 5.5 Binder of Arl2 (BART) Like 1

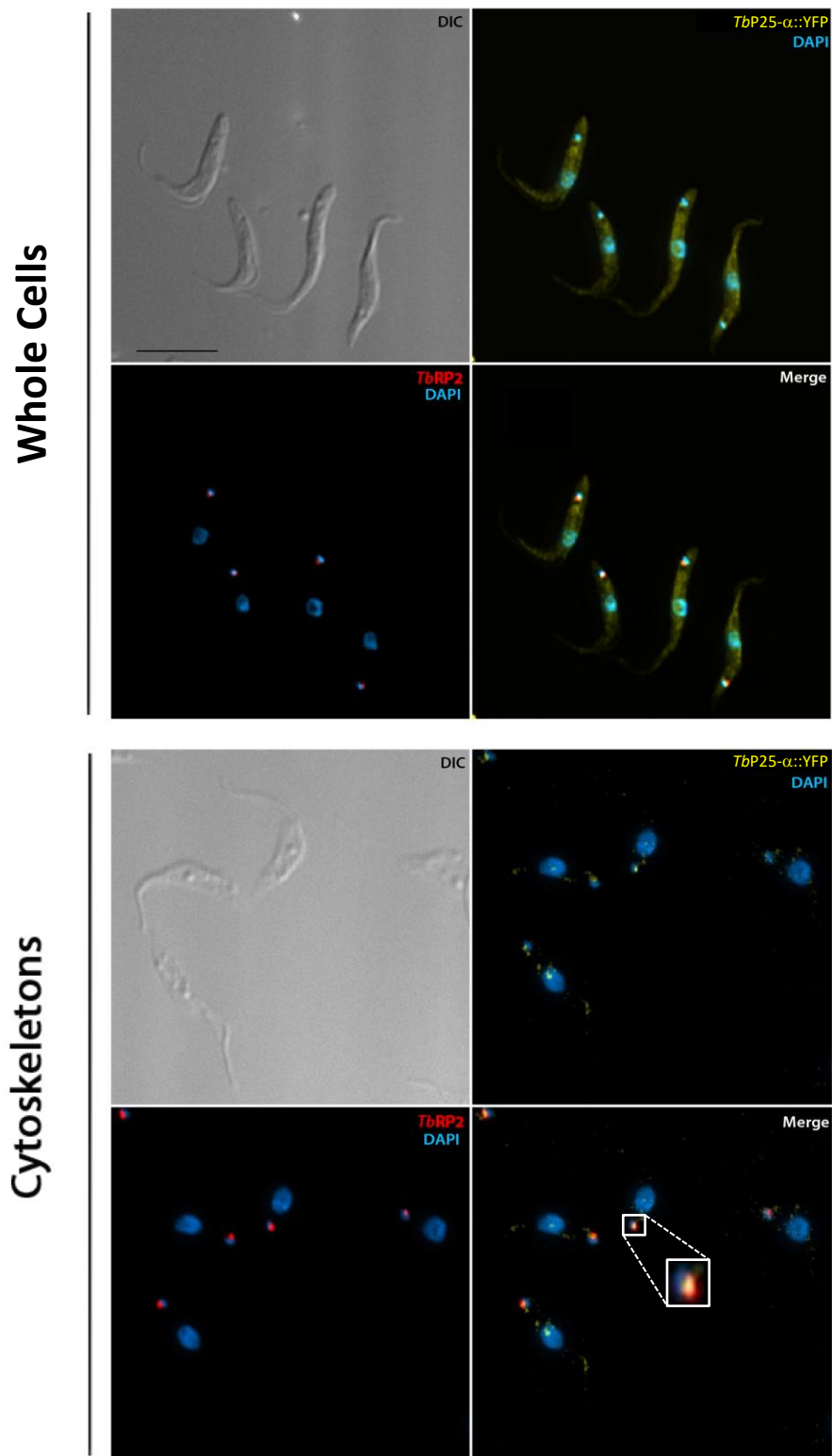
The localisation of *TbBARTL1* was examined in both whole cells and detergent-extracted cytoskeletons (Figure 5.7). In whole cells, *TbBARTL1::YFP* appears to have a uniform signal throughout the cytoplasm, but is excluded from the nucleus (Insert Panel A) and is also present within the flagellum. However, *TbBARTL1::YFP* was lost on detergent-extracted cytoskeleton cells. When *TbBARTL1::YFP* expressing procyclic cells were co-labelled with anti-*TbRP2* polyclonal antibody and only small number of cells appeared to have slightly elevated YFP signal at the base of flagellum (Insert Panel B) coincident with *TbRP2*. However, in majority of the cells are not.



**Figure 5.7 Immunofluorescence images showing the localisation of *TbBARTL1::YFP* in procyclic form *T. brucei*.** Both whole cells and cytoskeletal cells were co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Cells were visualised by DeltaVision deconvolution microscopy. Inset indicates that the YFP signal of *TbBARTL1* partially co-localised with *TbRP2* at the base of the flagellum. N indicates nucleus, which shows no *TbBARTL1::YFP* expression. Scale bar = 5  $\mu$ m.

## 5.6 Tubulin Polymerisation Promoting Protein (TPPP)/p25- $\alpha$

The intracellular localisation of *TbP25- $\alpha$ ::YFP* was also investigated in both whole cell and detergent-extracted cytoskeletons (Figure 5.8). In whole cells *TbP25- $\alpha$ ::YFP* appeared to be uniformly distributed throughout the cell body and flagellum, but in contrast to detergent extracted cytoskeletons *TbP25- $\alpha$ ::YFP* also appeared to be present in the nucleus of whole cells. It is difficult to conclude that any *TbP25- $\alpha$ ::YFP* signal is enriched at the base of the cilium, which co-localising with *TbRP2* antibody at basal body on whole cell sample. However, in detergent-extracted cytoskeleton cells, the *TbP25- $\alpha$ ::YFP* signal was largely lost, but, some signal remained including at the base of the flagellum co-localising with *TbRP2* (Figure 5.8; inset square). *TbP25- $\alpha$ ::YFP* was also observed along the line of the flagella and as a faint punctate intra-nuclear signal. These detergent-resistant signals were not investigated further.

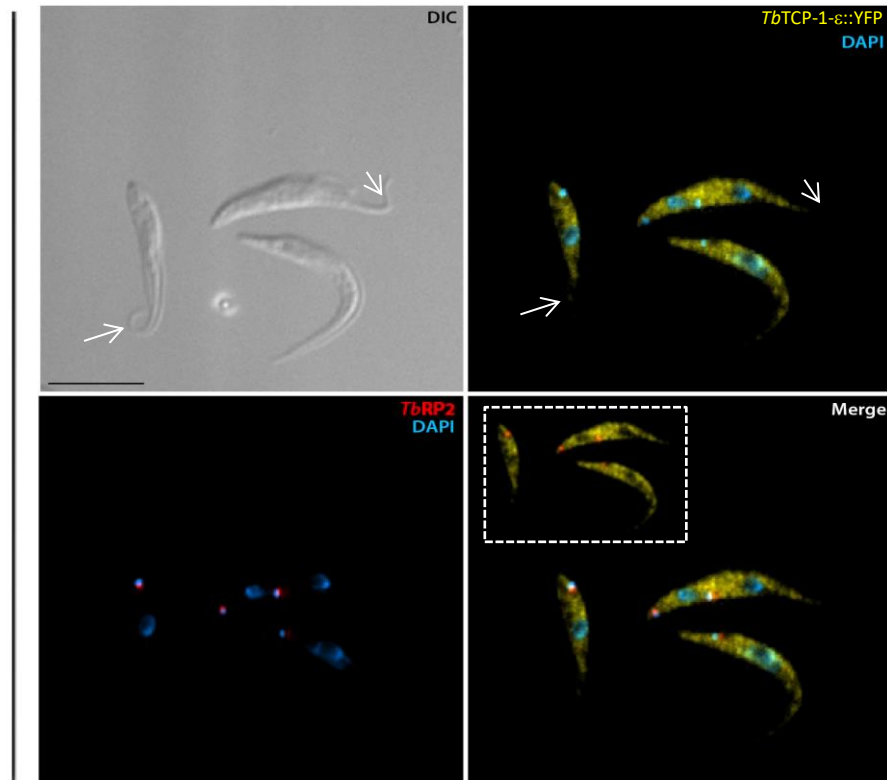


**Figure 5.8** Immunofluorescence images showing the localisation of *TbP25-α::YFP* in procyclic form *T. brucei*. Both whole cells and cytoskeletal cells were co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Cells were visualised by DeltaVision deconvolution microscopy. Scale bar = 5µm.

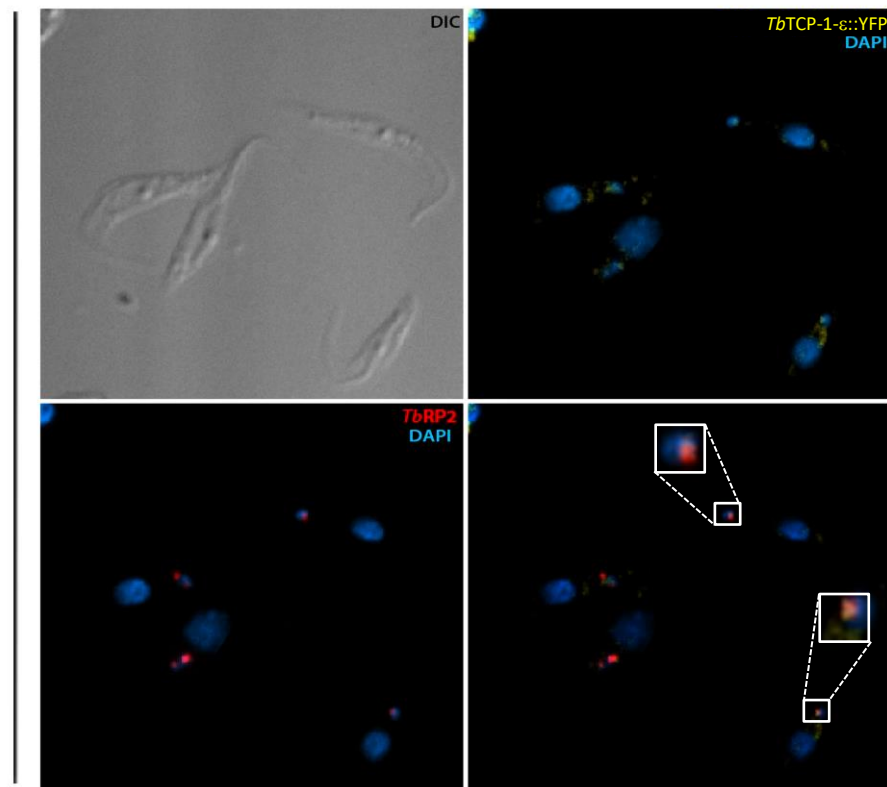
## 5.7 T-complex protein 1 epsilon (TCP-1-ε)

The intracellular localisation of *TbTCP1-ε::YFP* was investigated and revealed to be cytosolic, present within the cytoplasm, but was not observed in the nucleus (Figure 5.9; inset square) , nor the flagellum (Figure 5.9; white arrows). In detergent-extracted cytoskeletons, the *TbTCP1-ε::YFP* signal was largely absent. When the cells were co-labelled with the anti-*TbRP2* antibody, the foci of *TbTCP1-ε::YFP* at the base of the flagellum appeared to co-localise with *TbRP2* only in some (but not all) detergent-extracted cells.

## Whole Cells



## Cytoskeletons



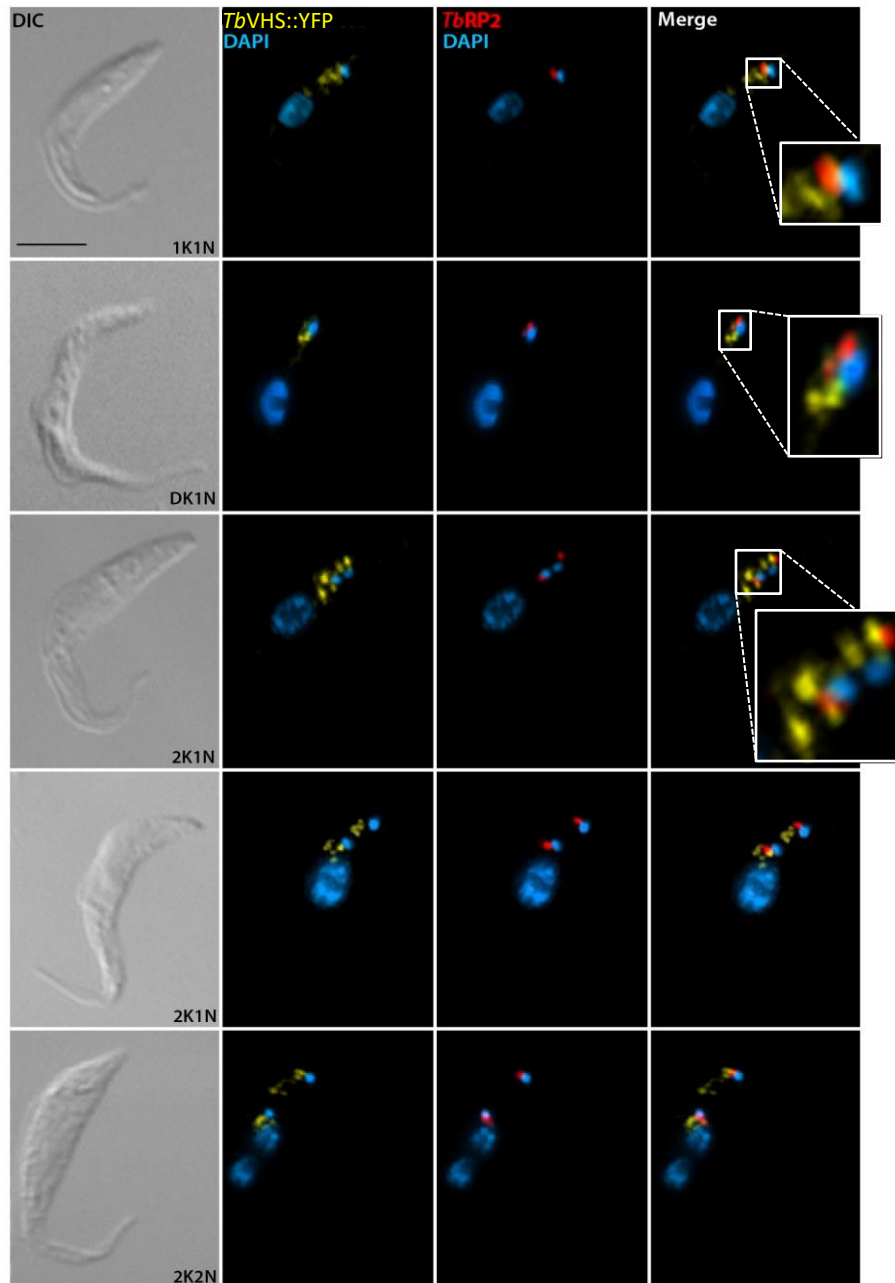
**Figure 5.9 Immunofluorescence images showing the localisation of *TbTCP-1-ε::YFP* in procyclic form *T. brucei*.** Both whole cells and cytoskeletal cells were co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Cells were visualised by DeltaVision deconvolution microscopy. White arrows/white square indicate that no expression of *TbTCP-1-ε::YFP* on flagella and nuclei respectively. Faint YFP signal of *TbTCP-1-ε* can be observed at the base of flagellum on cytoskeleton cells. Scale bar = 5µm.

## 5.8 VHS-domain containing protein

The localisation of *TbVHS::YFP* was determined in whole cells at various stages of the cell cycle (Figure 5.10). At beginning of the cell cycle (1K1N), *TbVHS::YFP* is observed in the vicinity of the kinetoplast, but the signal appears to be located distal to the kinetoplast in a region of the cell associated with the endocytic pathway. As the trypanosome cell progresses thorough the cell cycle (2K1N), the *TbVHS::YFP* signal duplicates and remains positioned distal to the two kinetoplasts throughout cell cycle. Co-labelling with the anti-*TbRP2* polyclonal antibody revealed little evidence of any co-localisation with *TbVHS::YFP* largely distal to *TbRP2* at the mature basal body.

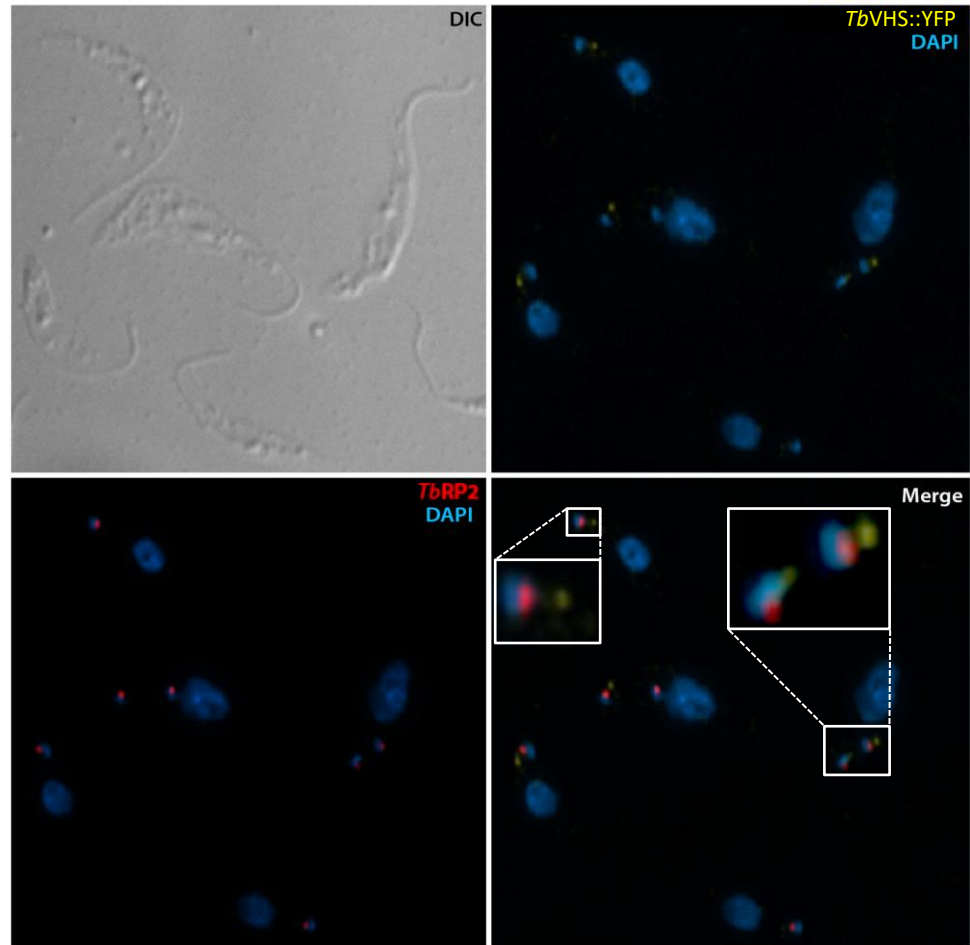
The localisation of *TbVHS::YFP* was then analysed in detergent-extracted cytoskeletons (Figure 5.11). A discrete focus of *TbVHS::YFP* signal was retained at a position close to the kinetoplast; this pattern was retained as the cell progressed thorough the cell division cycle. When cells were co-labelled with the anti-*TbRP2* antibody, the *TbVHS::YFP* signal was observed to be positioned more towards the anterior of the cell than *TbRP2* and the kinetoplast, thus there was no evidence of *TbVHS::YFP* co-localising with *TbRP2* at the mature basal body.





**Figure 5.10** Immunofluorescence images showing the localisation of *TbVHS::YFP* in procyclic form *T. brucei*. Whole cells were co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Cells were visualised by DeltaVision deconvolution microscopy. (Scale bar = 10µm)

## Cytoskeleton

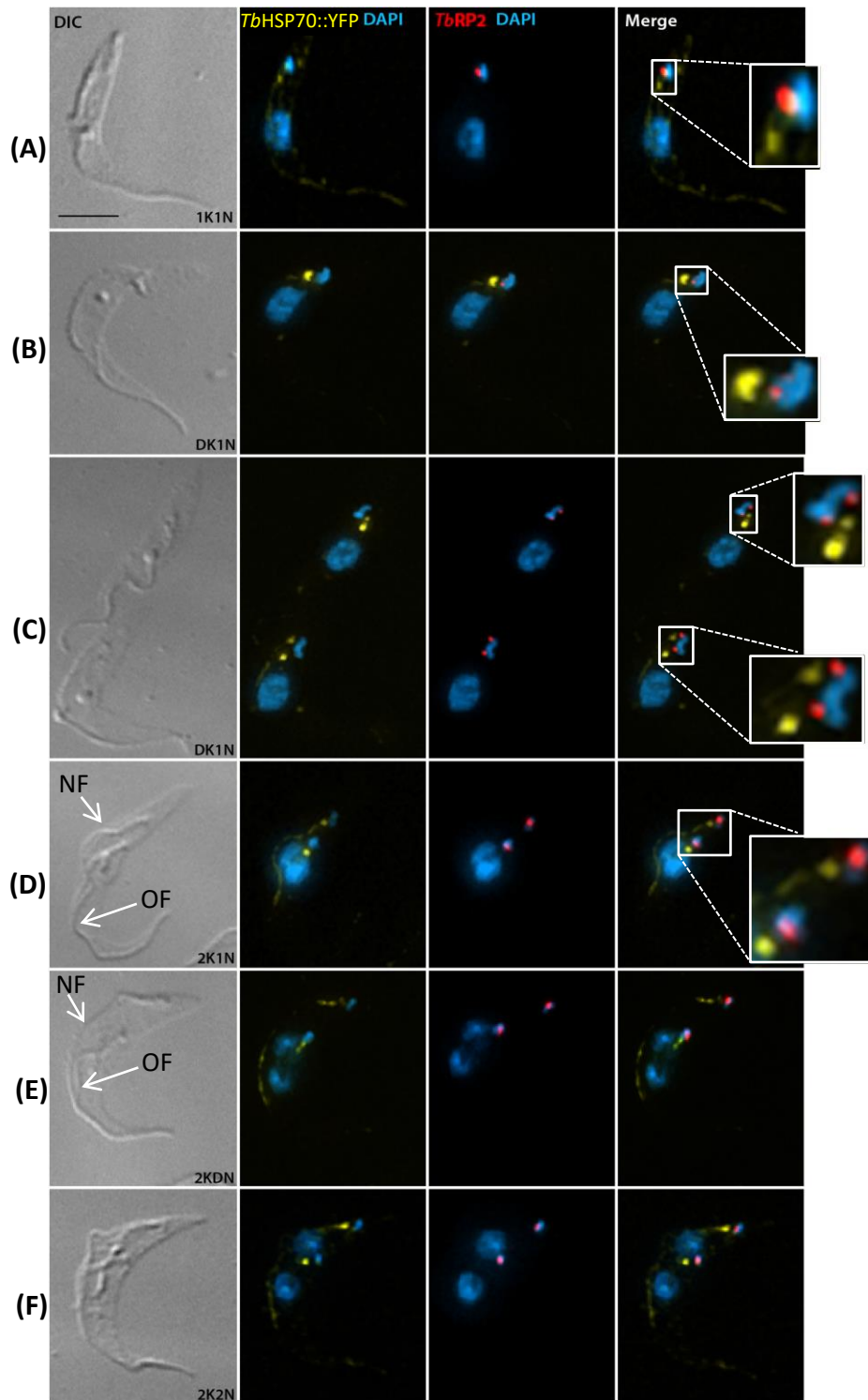


**Figure 5.11 Immunofluorescence images showing the localisation of *TbVHS::YFP* in procyclic form *T. brucei*.** Detergent extracted cytoskeletons were co-labelled with a *TbRP2* specific polyclonal antibody (red) and DAPI (blue). Cells were visualised by DeltaVision deconvolution microscopy. Scale bar = 5µm.

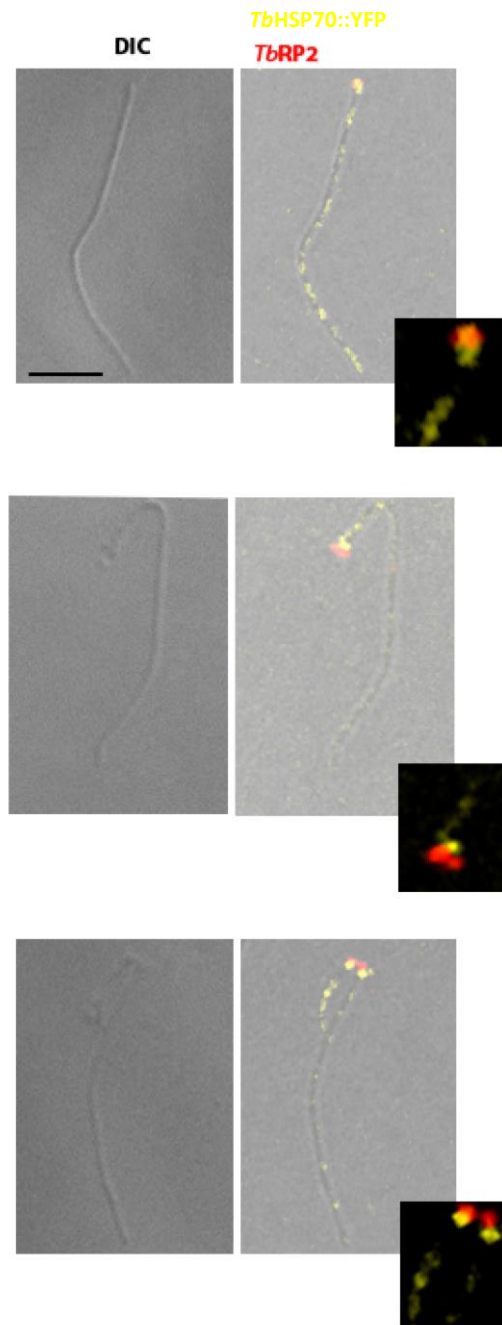
## 5.9 Heat Shock Protein 70 (HSP70)

The intracellular localisation of *TbHSP70::YFP* was investigated in detergent-extracted cytoskeletons. In these experiments, cells at various stages of the cell cycle were also co-labelled with anti-*TbRP2* antibody (Figure 5.12). In 1K1N cells, a discrete focus of *TbHSP70::YFP* was observed close to kinetoplast, but this signal did not seem to co-localise with *TbRP2*. However, a fainter *TbHSP70::YFP* focus was observed coincident with the *TbRP2* label (Figure 5.12; A-inset). A weak *TbHSP70::YFP* signal was also evident along the flagellum at this stage (Figure 5.12; A). As the cell progresses through the cell division cycle (Figure 5.12; B-C) before a new mature basal body has formed (this would be detected as a second *TbRP2* signal) the *TbHSP70::YFP* signal remained strong (Figure 5.12; B-inset). Later in the cell cycle, as evidenced by the detection of a second *TbRP2* positive mature basal body, *TbHSP70::YFP* signal closer to the new basal appeared to be weaker than the *TbHSP70::YFP* signal associated with the old basal body (Figure 5.12; C-inset). Later still, when the kinetoplasts were completely segregated, the intensity of *TbHSP70::YFP* signals distal to the old and new basal body (as detected by *TbRP2*), appeared to be more equivalent (Figure 5.12; D-inset). During mitosis, the *TbHSP70::YFP* signal seemed to be more elongated than at previous stages of the cell cycle. Throughout the cell cycle, *TbHSP70::YFP* also appeared to be localised in the flagellum, but it was noticeable that the *TbHSP70::YFP* signal appeared to be more intense in the new flagellum than the old (see white arrows Fig 5.12; D-E).

We then investigated whether *TbHSP70::YFP* was retained in flagella preparations by treating cells with a high salt buffer (0.5M NaCl) which preferentially depolymerises microtubules of the subpellicular cytoskeleton. The localisation of the *TbHSP70::YFP* was analysed in flagella preparations co-labelled with the anti-*TbRP2* antibody (Figure 5.13). In these flagella preparations *TbHSP70::YFP* signal was clearly evident both towards the flagellum base, and partially overlapping with *TbRP2*, but also along the length of the flagellum. The *TbHSP70::YFP* signal in the short new flagellum (Figure 5.13, Panel C) appeared particularly strong, which agreed with the labelling pattern observed in detergent-extracted cytoskeletons (Figure 5.12).



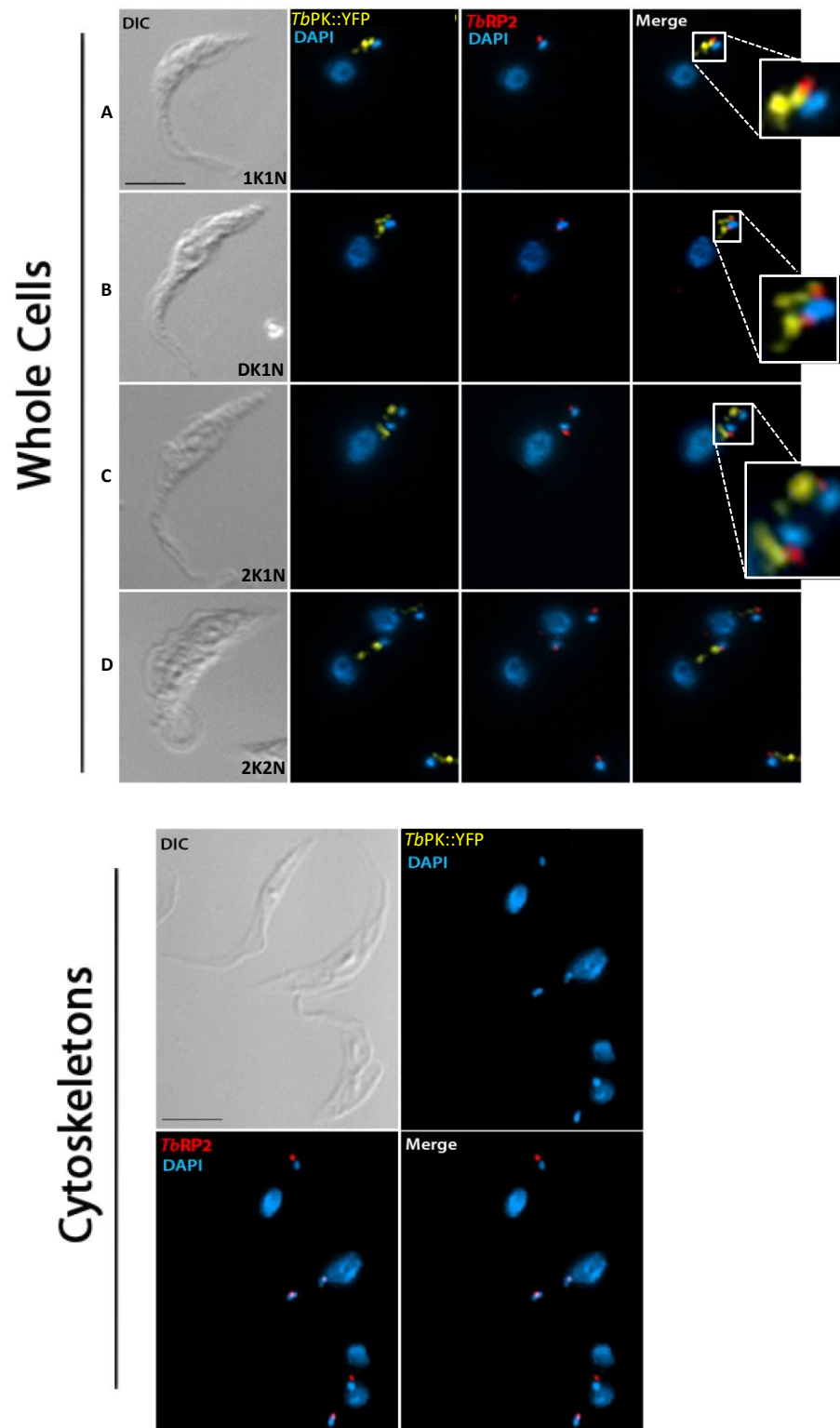
**Figure 5.12 Immunofluorescence images showing the localisation of *TbHSP70::YFP* in procyclic form *T. brucei*.** Cytoskeletal cells were co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Cells were visualised by DeltaVision deconvolution microscopy. OF (old flagellum); NF (new flagellum). Scale bar = 5 µm.



**Figure 5.13 Immunofluorescence images showing the localisation of *TbHSP70::YFP* in procyclic form *T. brucei*.** Cells were treated with 0.5M of NaCl and co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Flagella were derived from cells expressing *TbHSP70::YFP*, and visualised by DeltaVision deconvolution microscopy. Scale bar = 5μm.

## 5.10 Protein Kinase (PK) (*Tb927.9.6560*)

The intracellular localisation of *TbPK::YFP* was determined in whole cells at various stages of the cell cycle and shown to localise close to the kinetoplast, but not coincident with the mature basal body identified by *TbRP2* (Figure 5.14). The *TbPK::YFP* signal extends into the region between the kinetoplast and nucleus and as cells enter the cell cycle (i.e. kinetoplast undergoing replication and separation and new flagellum assembly has begun), a 'two-dot' pattern of *TbPK::YFP* can be seen distal to both the basal bodies identified by *TbRP2* (Fig 5.14-B). Once kinetoplast division is complete (2K1N), each kinetoplast remains associated with a focus of *TbPK::YFP* (Fig 5.14-C). In cells at the 2K2N stage of the cell cycle, the *TbPK::YFP* signal still appeared distal to the kinetoplasts and *TbRP2* (Fig 5.14-D). *TbPK::YFP* labelling was lost in detergent-extracted cytoskeletons (Figure 5.14-cytoskeletons).



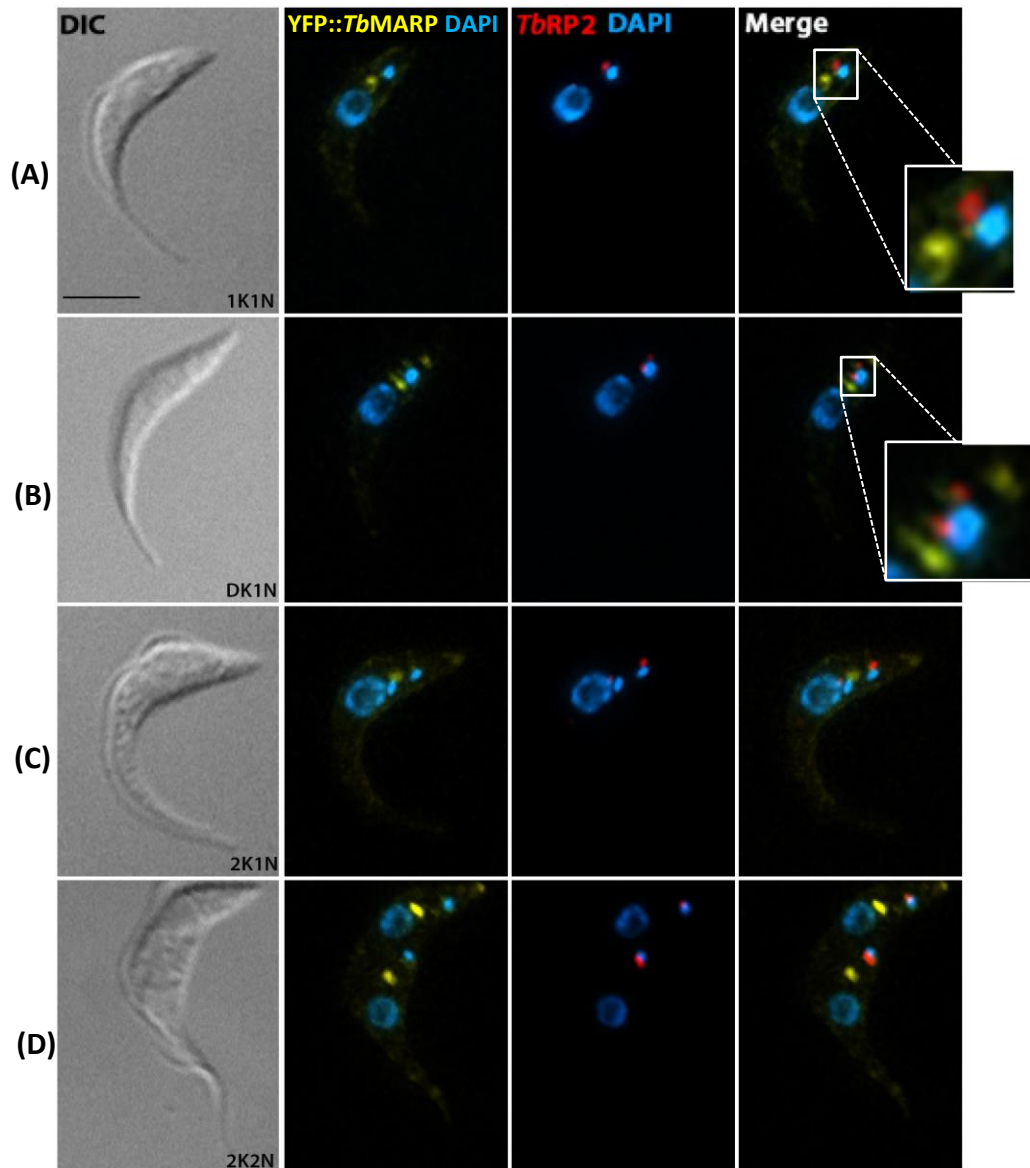
**Figure 5.14 Immunofluorescence images showing the localisation of *TbPK::YFP* in procyclic form *T. brucei*.** Both whole cells and cytoskeletons were co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Cells were visualised by DeltaVision deconvolution microscopy. Scale bar = 5  $\mu$ m.

## 5.11 Microtubule associated repetitive protein (MARP) (*Tb927.10.10360*)

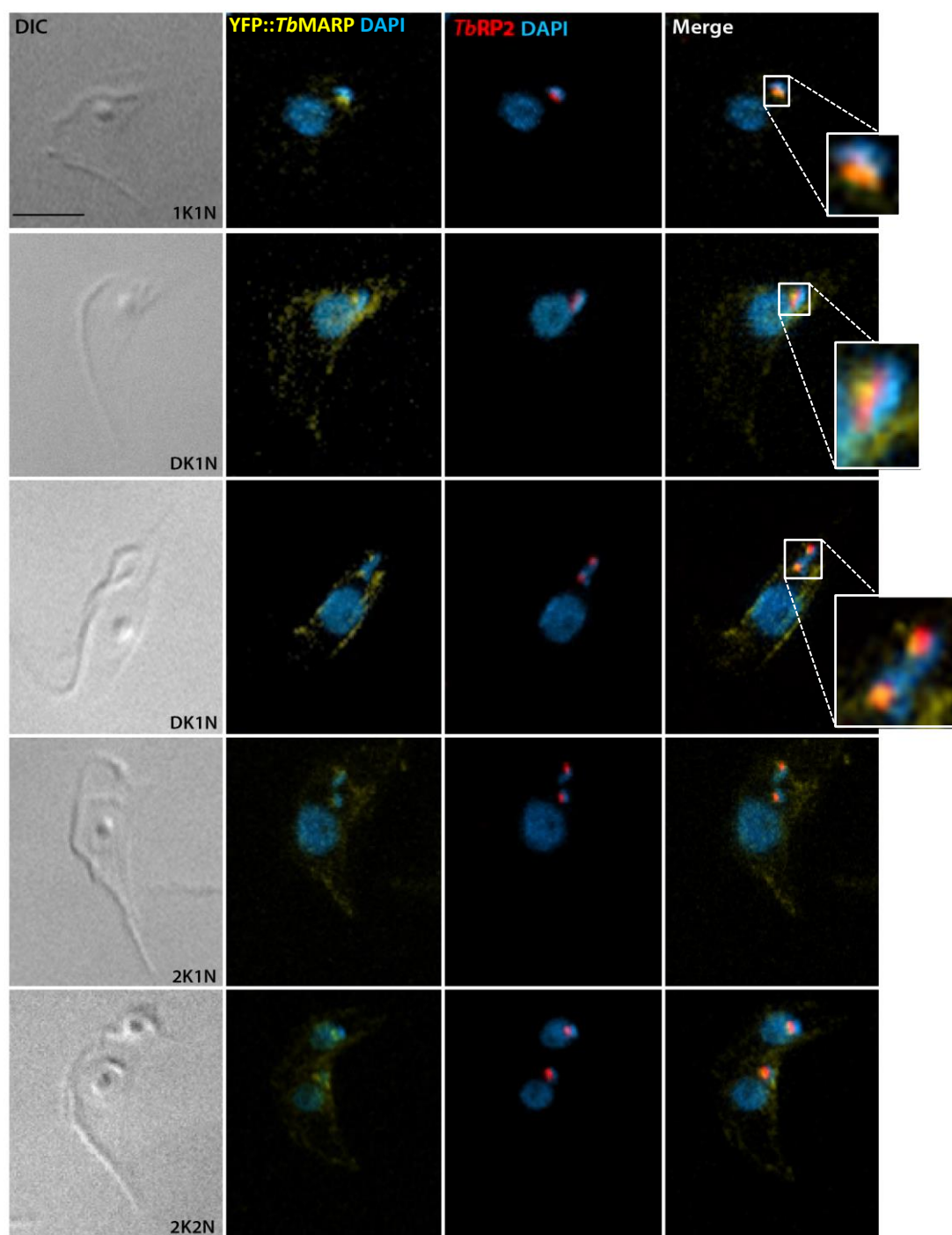
The localisation of YFP::*Tb*MARP was investigated in whole cells co-labelled with anti-*Tb*RP2 antibody (Figure 5.15). In cells at an early stage of the cell cycle (1K1N), a distinct and intense YFP::*Tb*MARP signal was observed between the nucleus and kinetoplast. YFP::*Tb*MARP also localised at: (i) the subpellicular microtubule cytoskeleton (present as a weak and dispersed signal), and (ii) adjacent to the kinetoplast/mature basal body (again as a weak signal) (Figure 5.15-A). From this imaging it was not apparent that YFP::*Tb*MARP co-localised with *Tb*RP2 at the mature basal body. When two *Tb*RP2 signals are present (i.e. after the probasal body has matured to a basal body and the kinetoplast is initiated replication), there is a single intense focus of YFP::*Tb*MARP distal to the basal bodies and kinetoplast. The basal body and subpellicular microtubule localisation of YFP::*Tb*MARP can also still be seen (Figure 5.15-B). When kinetoplasts are fully segregated, the intense YFP::*Tb*MARP signal seemed elongated (Figure 5.15-C). After mitosis is completed, two strong YFP::*Tb*MARP signals (between N and K) can be observed in addition to the weaker microtubule cytoskeleton and basal body proximal signals (Figure 5.15-D).

Detergent extraction resulting in the generation of *T. brucei* cytoskeletons revealed that the intense signal between nucleus and kinetoplast was lost, but that the basal body and microtubule cytoskeletal localisation of YFP::*Tb*MARP became more evident. Cytoskeletons were also co-labelled with the anti-*Tb*RP2 antibody, and indicated that YFP::*Tb*MARP at least partially co-localised with *Tb*RP2 (Figure 5.16). To investigate the basal body localisation in more detail, isolated flagellar preparations were made on cells expressing YFP::*Tb*MARP and co-labelled with the anti-*Tb*RP2 antibody. Visualisation of flagella from various stages of the cell division cycle revealed the YFP::*Tb*MARP signal was retained on isolated flagella and appeared to be located distal to the *Tb*RP2 signal i.e. further from the kinetoplast and more proximal to the axoneme (Figure 5.17).

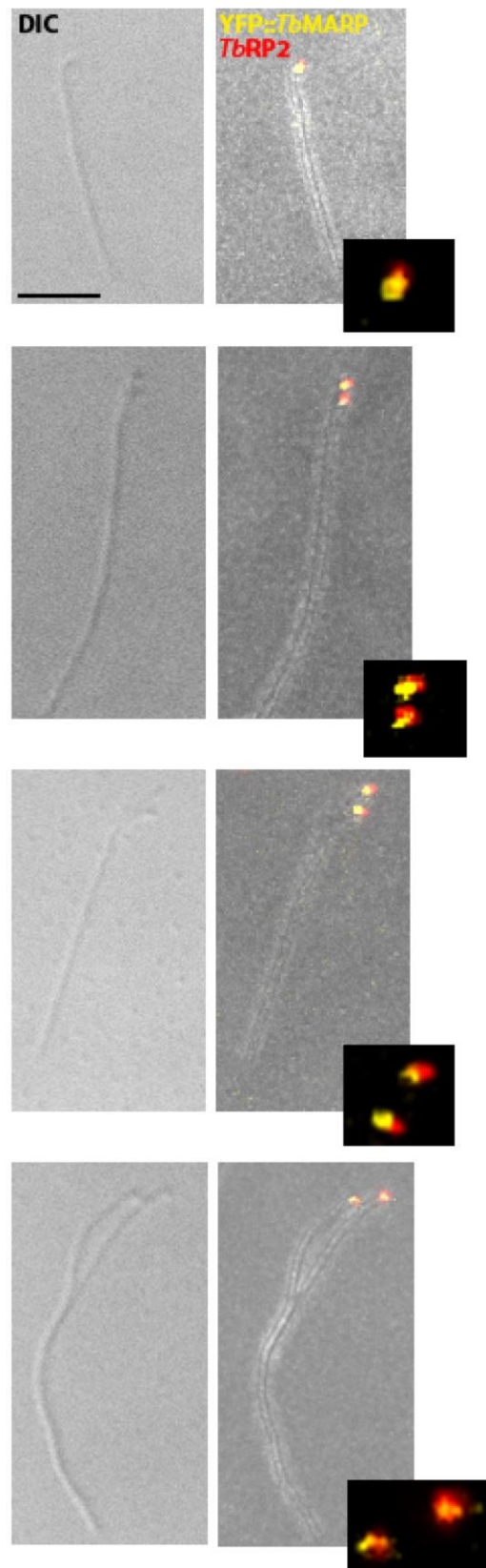




**Figure 5.15 Immunofluorescence images showing the localisation of *TbMARP::YFP* in procyclic form *T. brucei*.** Whole cells were co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Cells were visualised by DeltaVision deconvolution microscopy. Scale bar = 5  $\mu$ m.



**Figure 5.16 Immunofluorescence images showing the localisation of *TbMARP::YFP* in procyclic form *T. brucei*.** Detergent-extracted cytoskeletons were co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Cells were visualised by DeltaVision deconvolution microscopy. Scale bar = 5  $\mu$ m.

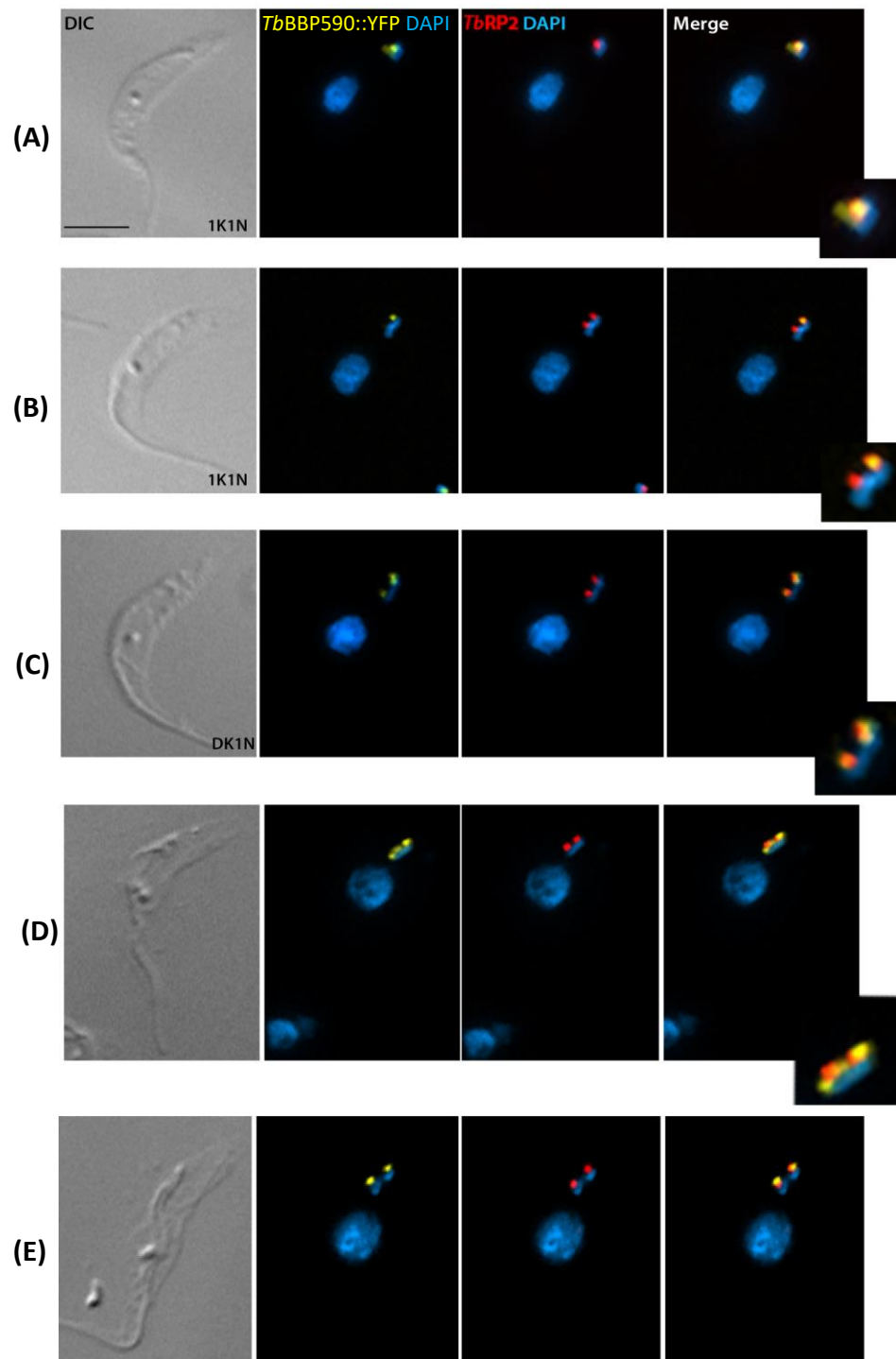


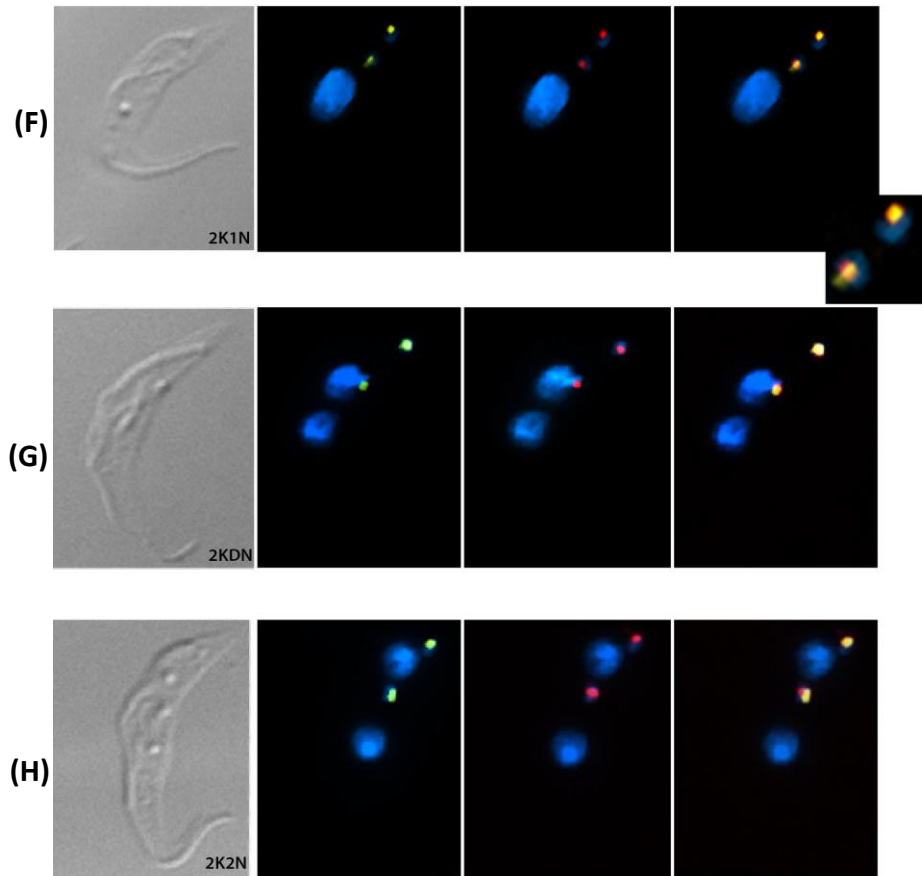
**Figure 5.17 Immunofluorescence images showing the localisation of YFP::TbMARP in procyclic form *T. brucei*.** Salt-extracted flagella were co-labelled with the TbRP2 antibody (red) specific for the mature basal body and DAPI (blue). Flagella were derived from cells were expressing YFP::TbMARP, and visualised by DeltaVision deconvolution microscopy. Scale bar = 5  $\mu$ m.

## 5.12 *TbBBP590*

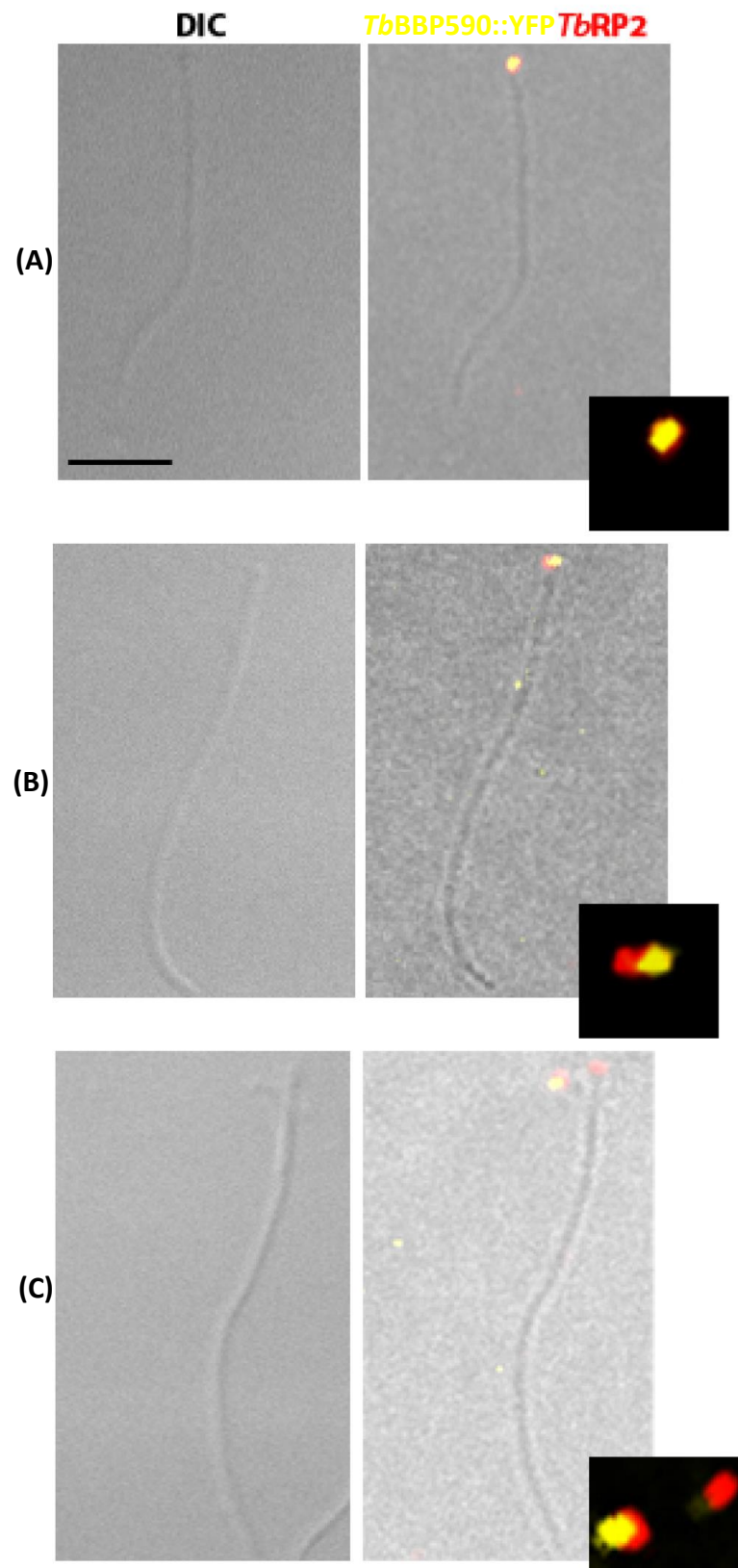
The localisation of *TbBBP590::YFP* was investigated in detergent-extracted cytoskeletons that were co-labelled with the anti-*TbRP2* antibody (Figure 5.18). During the early stage of the cell cycle (1K1N1F), a single discrete foci of *TbBBP590::YFP* was evident, and which overlapped with *TbRP2*. In addition, there was a rather weaker and diffuse *TbBBP590::YFP* signal adjacent to *TbRP2* (Figure 5.18-A). Interestingly, when the cell proceeds through the cell cycle, and two mature basal bodies clearly identified by the anti-*TbRP2* antibody, *TbBBP590::YFP* can only be detected at the newly formed mature basal body and not on the 'old' basal body (Figure 5.18-B). As the cell cycle progresses and the kinetoplast elongates prior to segregation, *TbBBP590::YFP* signal was reacquired on the older basal body (initially as a slightly fainter signal) (Figure 5.18-C). As the cell cycle progresses, and the kinetoplasts begin to segregate, the *TbBBP590::YFP* signal associated with the older basal body increases in intensity (Figure 5.18-D), and remains equal through the rest of the cell division cycle (Figure 5.18 D-H).

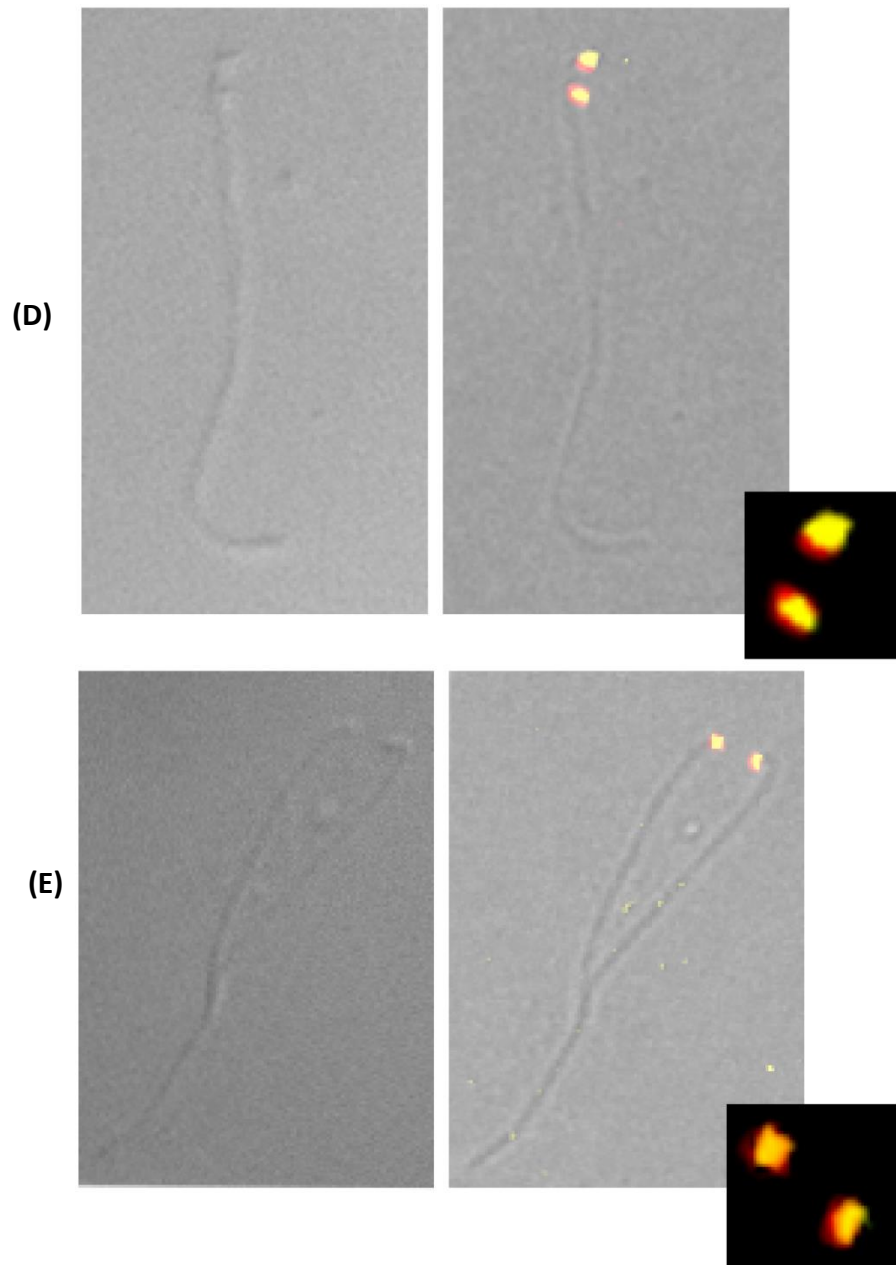
Expression of *TbBBP590::YFP* was also examined in isolated flagella. This confirmed the unusual pattern of *TbBBP590::YFP* expression shown in Figure 5.18; the detection of *TbRP2* was again used as marker for the mature basal body. From these images it is clearly evident that the *TbBBP590::YFP* signal associated with the basal body that nucleates the flagellum is initially present, but then lost as a new basal body is formed and the new flagellum nucleated. The 'new' basal body is now *TbBBP590::YFP* positive, and *TbBBP590::YFP* is only reacquired by the old basal body at a later time point (Figure 5.19; A-E).





**Figure 5.18 Immunofluorescence images showing the localisation of *TbBBP590::YFP* in procyclic form *T. brucei*.** Detergent-extracted cytoskeletons were co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Cells were visualised DeltaVision deconvolution microscopy. Scale bar = 5  $\mu$ m



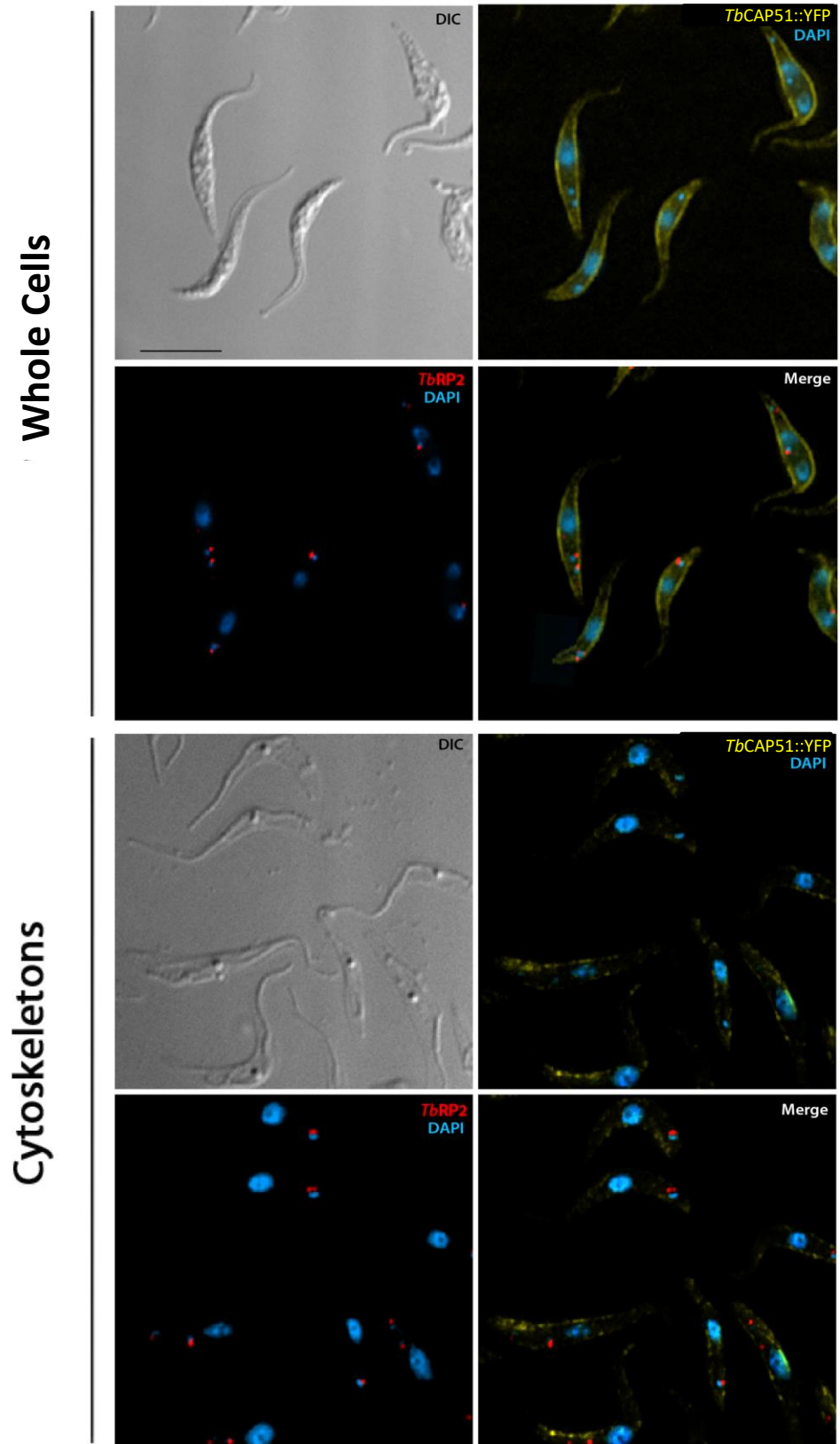


**Figure 5.19 Immunofluorescence images showing the localisation of *TbBBP590::YFP* in procyclic form *T. brucei*.** Salt-extracted flagella were co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Flagella were derived from cells were expressing *TbBBP590::YFP*, and visualised by DeltaVision deconvolution microscopy. Scale bar = 5  $\mu\text{m}$



### **5.13 TbCytoskeletal-associated protein (CAP51) (*Tb927.7.2640*)**

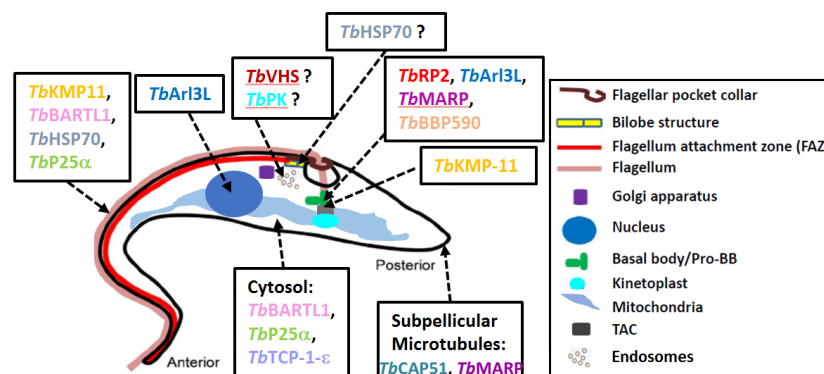
The localisation of *TbCAP51::YFP* was examined in both whole cells and in detergent-extracted cytoskeletons. In whole cells, *TbCAP51::YFP* clearly localised to the cell body with the signal stronger at the periphery indicating a subpellicular corset location. In detergent-extracted cells, *TbCAP51::YFP* was retained on the subpellicular microtubule cytoskeleton, but there were also punctate signals within the cell body. Both the whole cells and cytoskeletons were co-labelled with the anti-*TbRP2* antibody (to locate the mature basal body); there was no obvious co-localisation between *TbCAP51::YFP* and *TbRP2* (Fig 5.20).



**Figure 5.20** Immunofluorescence images showing the localisation of *TbCAP51::YFP* in procyclic form *T. brucei*. Whole cells and detergent-extracted cytoskeletons were co-labelled with the *TbRP2* antibody (red) specific for the mature basal body and DAPI (blue). Cells were visualised DeltaVision deconvolution microscopy. Scale bar = 5  $\mu\text{m}$ .

## 5.14 Summary

The localization of all eleven selected proteins in whole cells and/or detergent-extracted cells was determined (see summary Fig 5.21 and table 5.1). This was accomplished by pPOT system, which provides high-throughput and scalable endogenous tagging of target genes with C or N-terminal YFP fusion proteins (Dean et al. 2015). Among all eleven proteins, four proteins (*TbArl3L*, *TbMARP*, *TbBBP590* and *TbKMP-11*) localised near the basal body area. *TbVHS*, *Tbprotein Kinase (TbPK)* are appeared to localise to the endomembrane system. *TbHSP70* has cell body and flagellum localisation on whole cell. However, in cytoskeleton cells, the expression of the *TbHSP70::YFP* is more distal to *TbRP2*, possibly near bilobe area and is continuously distributed on the flagellum. *TbBARTL1*, *TbP25a* and *TbTCP-1-ε* all have cell body localisation, only *TbTCP-1-ε* does not show any expression in flagellum, but the other two also localised to the flagellum. There are some additional localisations of *TbKMP-11* have been identified in this study; rather than co-localisation with the BB, *TbKMP-11* expression appeared to have 2-dot coincidence with one *TbRP2* at the basal body, and with more proximal signal. Moreover, the expression on the flagellum is not evenly distributed, as the expression goes weaker after the end of the cell body, an elevated signal was also observed at the end of cell body. The flagellar connector localisation was also confirmed when cells were co-labelled with AB1 (Briggs et al. 2004). The only non-BB or cell body related protein is *TbCAP51*, which localised to the subpellicular microtubules. In the next chapter, the functionality of those proteins will be investigated by RNAi ablation.



**Figure 5.21 Schematic summarising localisations of putative *TbRP2*-interacting proteins and nearby neighbours.** Anti-*TbRP2* was used as a marker for the basal body labelled in red text.

<b>Proteins</b>	<b>Localisation(s)</b>
<b><i>TbArl3L</i></b>	Basal body; Nucleus
<b><i>TbBARTL1</i></b>	Cell body; Flagellum
<b>HSP70</b>	Bilobe area and consciously on the flagellum
<b><i>TbVHS</i></b>	Endomembrane system
<b><i>TbP25-α</i></b>	Cell body; Flagellum
<b><i>TbTCP-1-ε</i></b>	Cell body
<b><i>TbMARF</i></b>	Basal body; subpellicular microtubules
<b><i>TbBBP590</i></b>	Basal body
<b><i>TbProtein Kinase</i></b>	Endomembrane system
<b><i>TbKMP-11</i></b>	Basal body; Flagellum; Flagellar connector; End of cell body
<b><i>TbCAP51</i></b>	Subpellicular microtubules

**Table 5.1** Table summarising the localisation(s) of selected putative *TbRP2*-interacting/near neighbour proteins.

## Chapter 6 Further characterisation of *TbRP2*-proximal proteins

### 6.1 Introduction

In this chapter, the putative function(s) of selected *TbRP2* interacting/proximal proteins are investigated using RNAi-mediated ablation. The effect of gene specific RNAi on growth and morphology of procyclic *T. brucei* cells was assessed, along with the effect on flagellum biogenesis and the expression and/or localisation of *TbRP2*. Additionally, for a selected subset of these putative *TbRP2* interacting/proximal proteins, the effect of *TbRP2* RNAi ablation on the expression and/or localisation of the protein of interest was investigated.

To carry out these studies, DNA encoding the open readings frames of interest (or defined sub regions thereof) were amplified by PCR and the resulting amplicons cloned into pGEMT© for nucleotide sequence validation. Following confirmation of wild type sequence (See Appendix III), DNA fragments were excised from pGEMT and sub-cloned into the vector p2T7-177 (Wickstead et al., 2002). Cell lines were established in smOX P9 cells (Poon et al., 2012), which allowed for inducible induction of RNAi, in cell lines that had been previously engineered to express YFP-tagged versions of the particular proteins of interest. This allowed for the efficiency of RNAi-mediated protein ablation to be assessed by immunoblotting and fluorescence microscopy approaches.

### 6.2 *TbKMP-11*

*TbKMP-11* was identified on multiple occasions in the BioID/SILAC experiments reported here, suggesting this protein is likely to be a genuine *TbRP2*-interacting/proximal protein. However, *TbKMP-11* has a complex intracellular localisation pattern in *T. brucei* (See Chapter 5.4), localising near to the basal body and

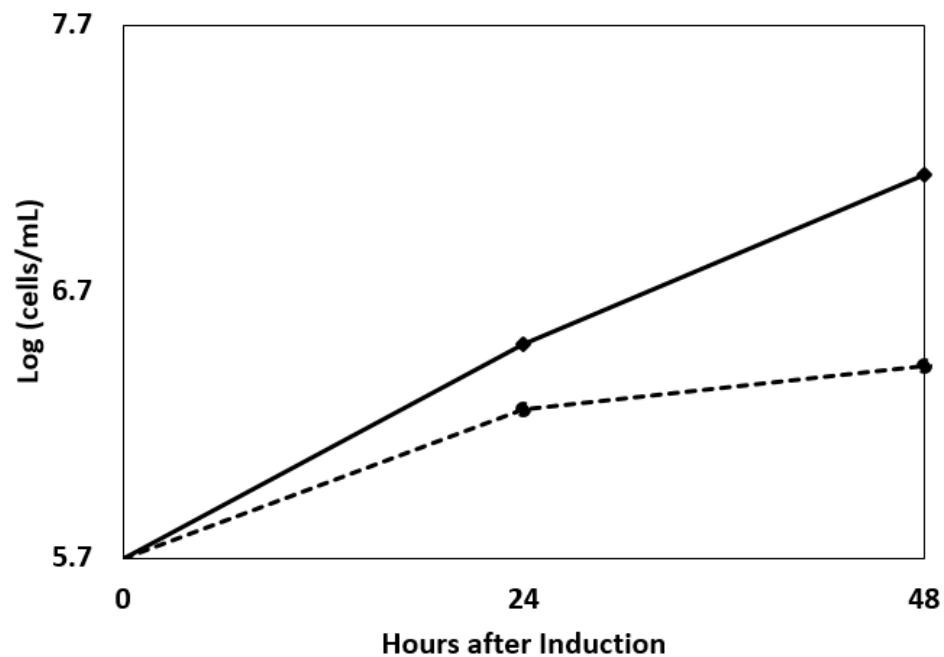
within the flagellum compartment itself. Thus, understanding where (but also why) *TbKMP-11* comes into close proximity to *TbRP2* is a challenge. For instance, is *TbKMP-11* juxtaposed to *TbRP2* at the basal body, or does it come into contact with *TbRP2* as it passes *TbRP2* located at the transitional fibres which demarcate the dynamic boundary between the cell body and flagellum i.e. the 'ciliary gate'.

Although the effect that *TbKMP-11* RNAi has on *T. brucei* cells has been investigated previously (and shown to affect basal body segregation and cytokinesis in both procyclic and bloodstream form cells; (Li and Wang, 2008)), *TbKMP-11* RNAi induction was repeated here to investigate the effect that *TbKMP-11* depletion had on *TbRP2* expression and basal body localisation. Preliminary experiments confirmed *TbKMP-11* RNAi resulted in the phenotypic defects previously described (Li and Wang, 2008). Following induction of *TbKMP-11* RNAi, a dramatic reduction in cell growth was evident within 24 hours (Fig 6.1-A). Analysis of cell growth was terminated after 48 hours of RNAi induction, as the induced cells became too 'monstrous' to allow a meaningful cell counts to be undertaken after this time. To confirm the efficiency of *TbKMP-11* RNAi knockdown, *TbKMP-11::YFP* expression was monitored by probing immunoblots of samples collected 24 and 48 hours post induction with an anti-GFP antibody (Fig 6.1-B). These immunoblotting experiments demonstrated that 24 hours post RNAi induction, *TbKMP-11::YFP* expression was dramatically reduced compared to non-induced controls. After 48 hours of induction, *TbKMP-11::YFP* expression was further reduced, although low levels of protein were still detectable. The anti- $\beta$ -tubulin antibody KMX-1 was used as a loading control to confirm equal loading from non-induced and RNAi induced cell populations (and in all subsequent RNAi induction experiments).

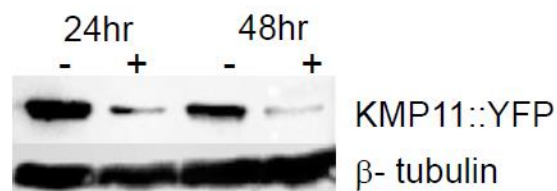
Cells from induced and non-induced were also prepared for immunofluorescence microscopy (Fig 6.1-C (1-4)). In non-induced cells (Fig 6.1-C-(1)), *TbKMP-11* localisation was identical to that reported in Chapter 5.4, but following RNAi induction the basal body and flagellar *TbKMP-11::YFP* signal was significantly reduced, and the reduction of the YFP signal is more obvious on flagella, as *KMP-11* localisation near the

kinetoplast is still detectable (Fig 6.1-C (2-4)). From these experiments, it was also evident that 48h after induction, *TbKMP-11* RNAi induced cells exhibited numerous cytological problems; (i) kinetoplast and basal body positioning/segregation defects, (ii) detached flagella, and (iii) cells possessed either multiple nuclei or were anucleate zoids (i.e. 1K0N). As these cells were co-labelled with an anti-*TbRP2* antibody, the loss of *TbKMP-11* on the basal body targeting and tethering of *TbRP2* was also evaluated. Despite the severe phenotypic abnormalities resulting from loss of *TbKMP-11*, expression and localisation of *TbRP2* at the basal body was clearly unaffected (Fig 6.1-C (3) as an example).

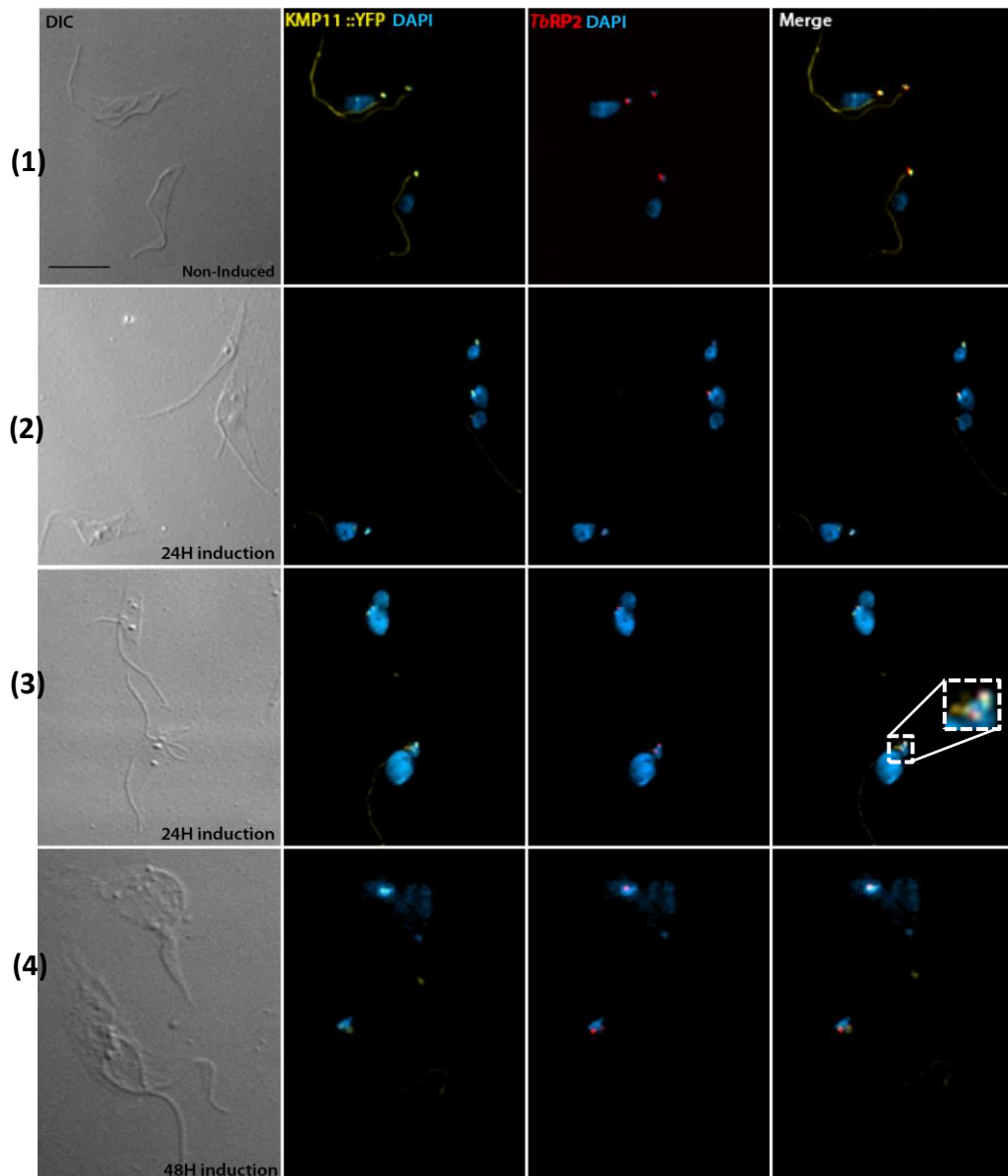
**(A)**



**(B)**



**(C)**



**Figure 6.1 RNA interference-mediated ablation of *TbKMP-11* results in multiple defects in procyclic form cells. (A)** Graph showing growth of *TbKMP-11*-RNAi induced (black dotted line) and non-induced (black solid line) procyclic trypanosome cells. **(B)** The *TbKMP-11::YFP* expression level was monitored by western blotting. Non-induced and induced cells were harvested after 24h and 48 h induction. KMX1 detection of  $\beta$ -tubulin was included as a loading control. **(C)** Non-induced cells at 24h/48h post induction *TbKMP-11*-RNAi cell line were harvested for immunofluorescence analysis. Morphological abnormalities in *TbKMP-11* RNAi cells can be observed after 24h induction, including kinetoplast/basal body segregation defects, flagella detachment and unequal segregation of nuclear DNA. Non-induced and RNAi cells were stained with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m.



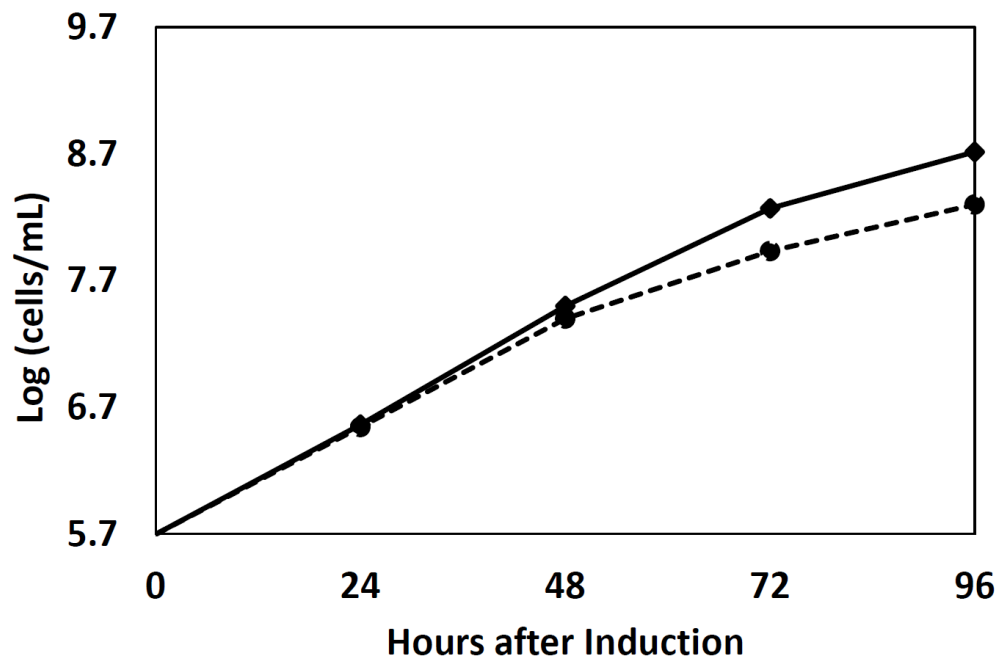
### 6.2.1 *TbKMP-11* flagellum localisation – relationship to FLAM3

The trypanosome flagellum emerges from the flagellar pocket and remains attached to the cell body for most of its length via the flagellum attachment zone (FAZ), but then extends a short distance beyond the anterior pole of cell as a ‘free flagellum’ (Ralston et al., 2009). Detailed analysis of *TbKMP-11* localisation indicated the *TbKMP-11::YFP* signal was more intense where the flagellum was attached to the cell body, and became weaker where the flagellum extended beyond the anterior pole of the cell. This pattern was strikingly similar to that observed for a protein designated FLAM3; which is reported to be a component of the FAZ (Sunter et al., 2015). In PCF cells, RNAi ablation of FLAM3 leads to a short new FAZ and results in a dramatic change in cell morphology with the generation of epimastigote-like cells (Sunter et al., 2015). The typical features of an epimastigote-form cell is a long free flagellum (i.e. the flagellum attached to the cell body is shortened) and the kinetoplast is juxtaposed to the nucleus toward the anterior end of the cell body.

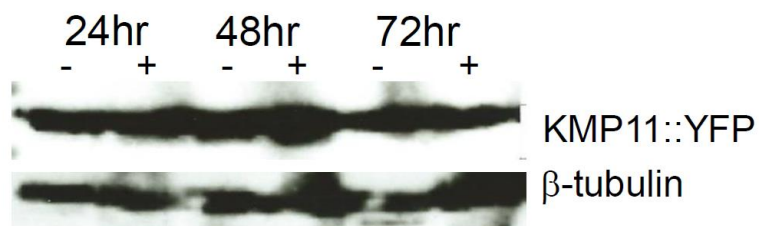
To investigate the relationship between *TbKMP-11* and FLAM3, cell lines were generated that allowed for inducible ablation of FLAM3 in procyclic cells expressing *TbKMP-11::YFP*. Although a growth defect was evident 24 hour after induction of FLAM3 RNAi (Fig 6.2-A), immunoblotting revealed that *TbKMP-11* expression was unaffected (Fig 6.2-B). Although knockdown of FLAM3 expression was not directly ascertained in these induction experiments the phenotypic consequence of FLAM3 RNAi was examined by immunofluorescence microscopy (Fig 6.2-C-(1-4)). FLAM3 RNAi induced cells clearly displayed the previously reported FLAM3 RNAi phenotype as early as 24 hours post induction; i.e. the majority of cells were undergoing morphological change from trypomastigote to epimastigote-like cells (Fig 6.2-C-(2); white square). Following longer periods of RNAi induction (48h and 72h) cells were observed in which both the kinetoplast and nuclei had duplicated and segregated, but these cells were unable to undergo cytokinesis resulting in the generation of multi-nucleate cells (Fig 6.2-C-(3-4)). However, even in these grossly abnormal cells, the localisation of *TbKMP-11::YFP* within the flagellum was unaffected; i.e. *TbKMP-11::YFP* signal was strong where the flagellum was attached to cell body and of diminished

intensity in the 'free flagellum'. The interpretation of this result is discussed in the summary section at the end of this chapter.

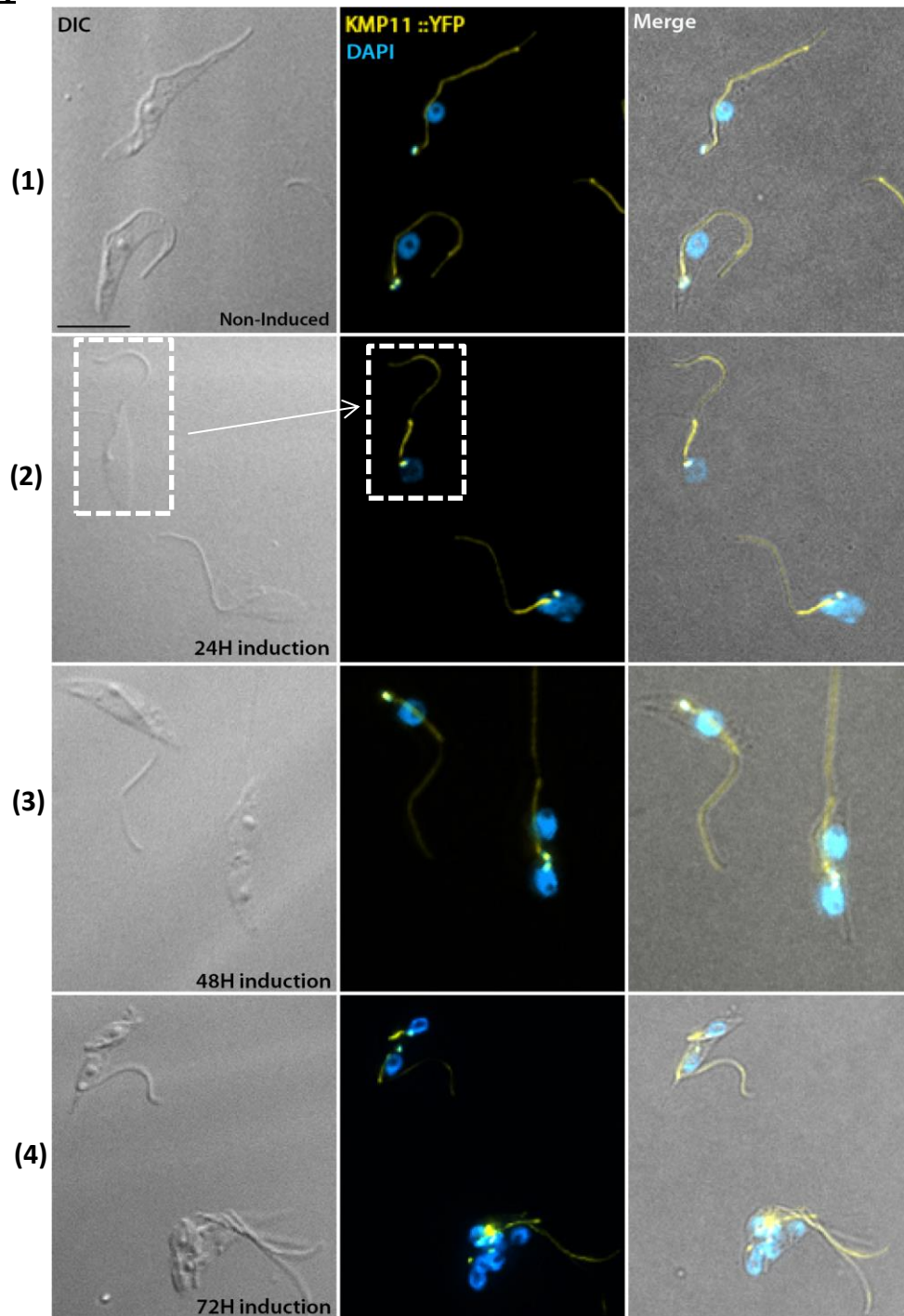
**(A)**



**(B)**



**(C)**



**Figure. 6.2 The expression and localisation of *TbKMP-11* is not dependent upon FLAM 3.** Cell lines capable of FLAM 3 inducible RNAi and expressing *TbKMP-11::YFP* were generated. **(A)** Growth curves of *TbKMP-11::YFP*, FLAM 3 RNAi - smOXP9 cells – RNAi induced (broken black line) and non-induced (solid black line). **(B)** The expression of *TbKMP-11::YFP* level on FLAM3 RNAi background was analysed by immunoblotting with anti-GFP antibody; KMX-1 acting as loading control. **(C)** FLAM3 RNAi causes a dramatic change in cell morphology, but the expression level and localisation of *TbKMP-11* (yellow) in detergent extracted cytoskeletons remains unaffected. The box designated with broken white line indicates an epimastigote-like cell formed after FLAM 3 RNAi induction. Scale bars = 10  $\mu$ m.

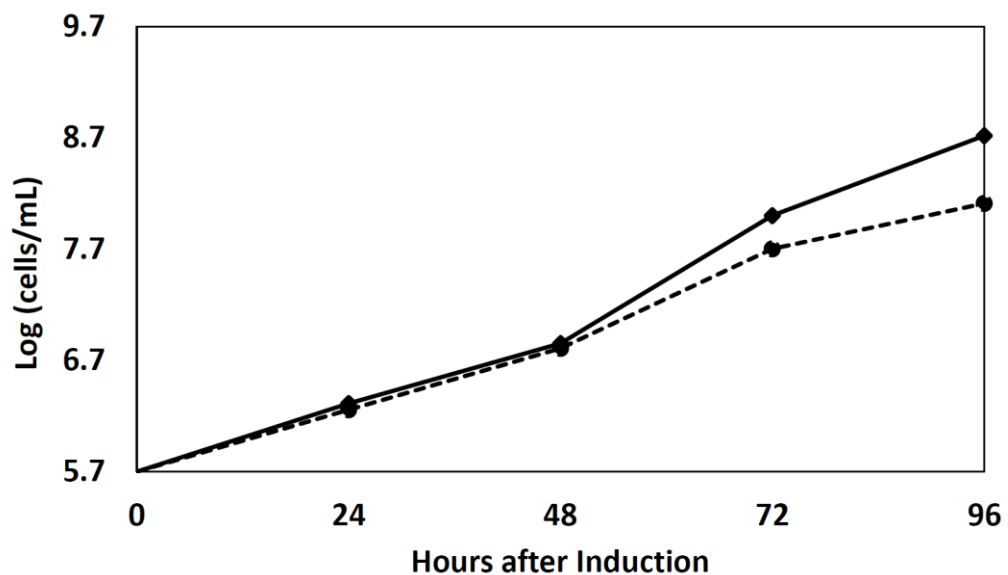
### 6.2.2 *TbKMP-11* flagellum localisation – relationship to calmodulin

In *T. brucei*, the  $\text{Ca}^{2+}$  binding protein calmodulin (CaM) is present within all three structural domains of the PFR, as well the linkage between PFR and axoneme, and PFR to FAZ (Ridgley et al., 2000). Based on the localisation study of *TbKMP-11::YFP* in chapter 5.4, the *TbKMP-11::YFP* signal extends beyond the anti-PFR2 antibody (L8C4) at proximal end of the flagellum, suggesting *TbKMP-11* is more likely to be associated with the axoneme than the PFR. To investigate the effect of CaM-depletion on *TbKMP-11* expression and localisation, cell lines were established to allow CaM-RNAi ablation on cells expressing *TbKMP-11::YFP*. Following CaM RNAi induction, cell density was determined in induced and non-induced populations of cells (Fig 6.3-A) and a pronounced growth defect was observed 24 hours after induction of RNAi. Expression of *TbKMP-11::YFP* was investigated by immunoblotting and revealed to be unaffected by CaM depletion until 48 hours post induction, but the expression was slightly reduced at 72 hours post induction (Fig 6.3-B). Specific detection of CaM was not possible in this experiment (due to the lack of available CaM-specific antibodies), but the flagellum detachment observed following induction of CaM RNAi was in accordance with the phenotypic effect previously published for this protein (Ginger et al., 2013).

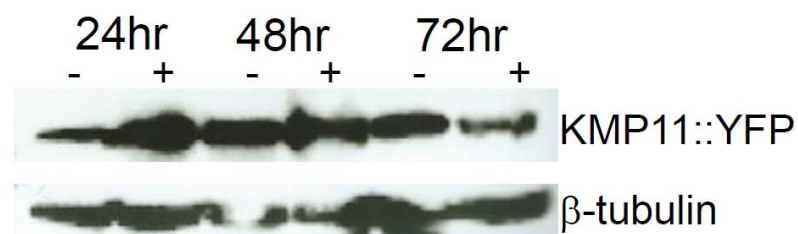
The localisation and expression of *TbKMP-11::YFP* was investigated in CaM-RNAi induced cells by immunofluorescence microscopy; with cells co-labelled with anti-*TbRP2* antibody to demarcate the mature basal body (Fig 6.3-C-(1-5)). In the non-induced cells, the localisations of *TbKMP-11::YFP* appeared as described earlier (Fig 6.3-C-(1)). After 24 hours of induction, *TbKMP-11::YFP* expression and localisation to near the kinetoplast seemed unaffected. However, short cells with an abnormally long free flagellum began to appear in the cell population (Fig 5.3-C-(2-4)). The depletion of CaM is reported to cause a PFR assembly defect, although the structure of the axoneme remains unaffected (Ginger et al., 2013). However, as the loss of CaM disrupts connection between the axoneme and the PFR (and thus to the cell body via the FAZ); cells with detached flagellum arise following CaM RNAi ablation (Fig 6.3-C-

(3)). At later time points following RNAi induction, cells with grossly abnormal morphologies and multinucleate cells were observed (Fig 6.3-C-(5)). Interestingly, in the CaM RNAi cells where the flagellum is clearly detached from the cell body, a foci of *TbKMP-11::YFP* is observed at the end of cell body (white rectangular box), along with a 'blob-like' signal at the distal tip of the flagellum (yellow rectangular box). In cells where the flagellum was still evidently attached, *TbKMP-11::YFP* signal was only detected within the flagellum attached to the cell body, and not in the free flagellum beyond the anterior pole.

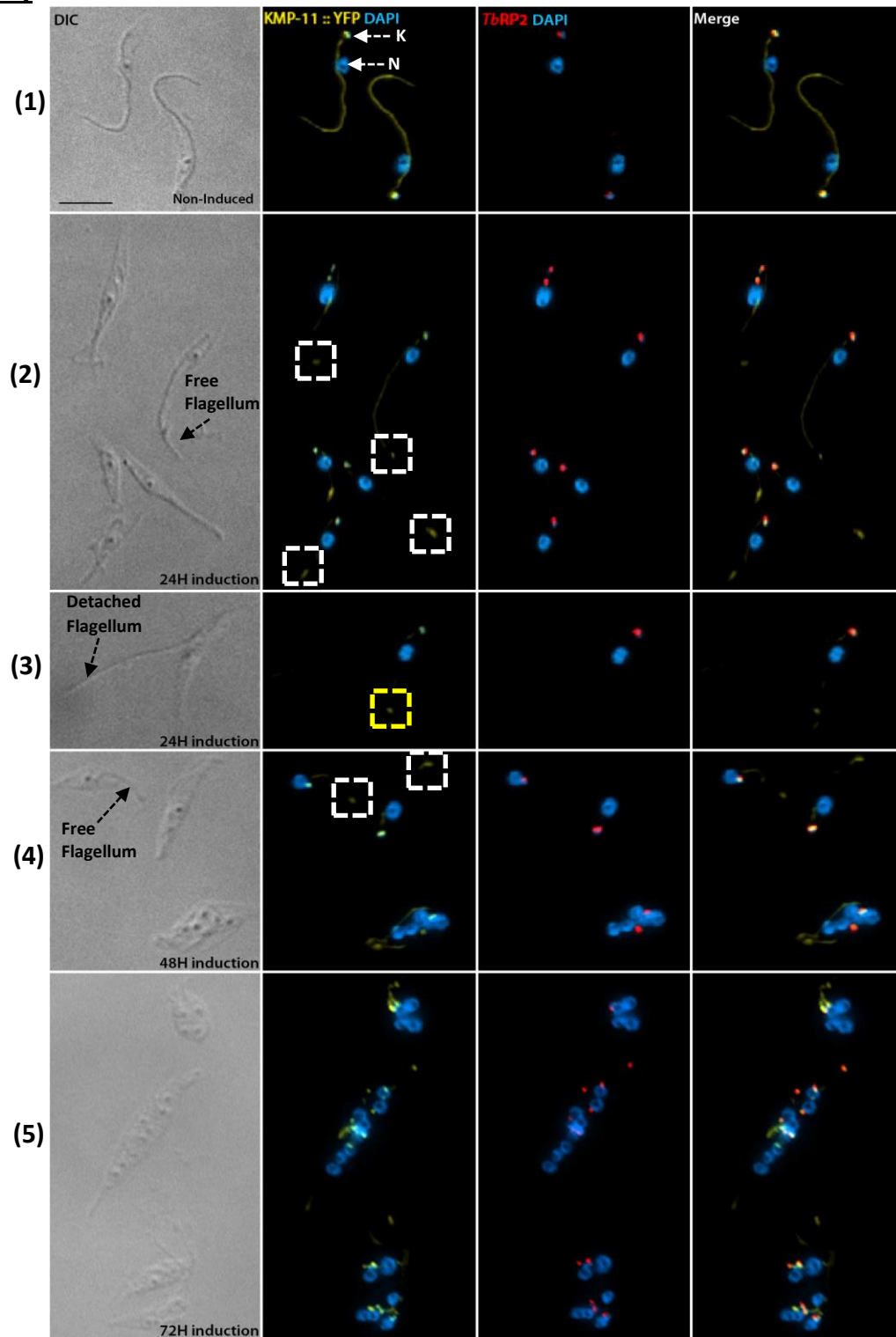
**(A)**



**(B)**



(C)



**Figure 6.3 The level of *TbKMP-11::YFP* expression was examined following CaM RNAi induction. (A) Growth curve of the *TbKMP-11::YFP*-CaM RNAi cells (Non-induced - solid black line; RNAi Induced - broken black line). (B) Immunoblot analysis of *TbKMP-11::YFP* expression following induction of CaM RNAi with anti-GFP antibody. (C) Immunofluorescence analysis of *TbKMP-11* expression and localisation. In cells with detached flagella, or flagella, which extended beyond the cell body, *TbKMP-11* YFP could not be detected. Both non-induced and induced cells were labelled by anti-*TbRP2* antibody (red) and DAPI (Blue). Scale bars = 5  $\mu$ m.**

### 6.3 *Tb*Cytoskeletal-associated protein (CAP51)

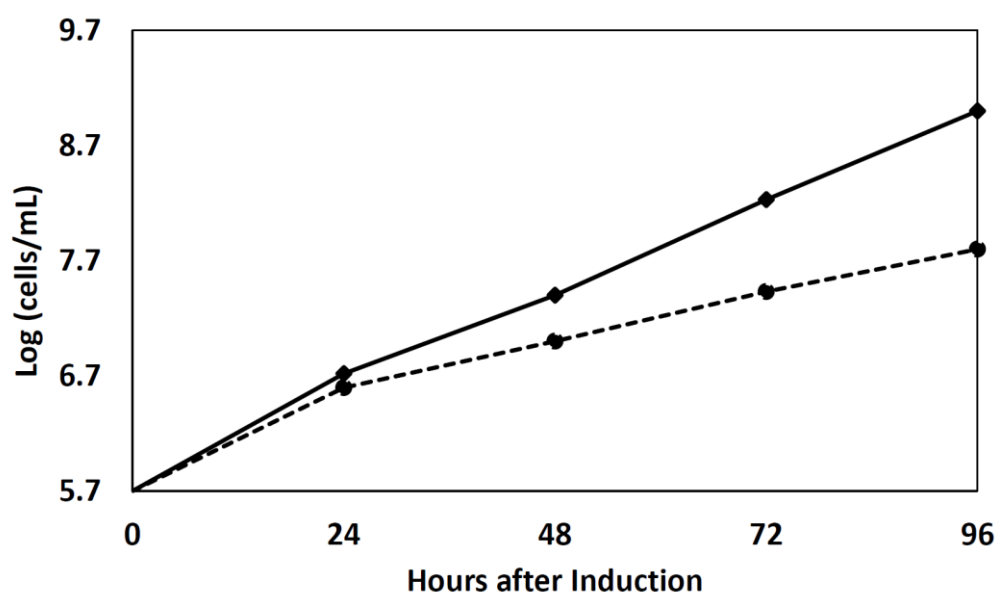
Identification of the *Tb*CAP51 encoded by *Tb*927.7.2640 as a *Tb*RP2-interacting/near neighbour protein is difficult to reconcile due to its localisation on the SPMT cytoskeleton (see Chapter 5.13). Nevertheless, the effect of *Tb*CAP51 RNAi induction on cell growth, cell morphology and flagellum biogenesis was investigated, along with its effect on the localisation of *Tb*RP2 at the mature basal body. *Tb*CAP51 RNAi was induced and cell growth monitored every 24 hours (Fig 6.5-A). RNAi mediated ablation of *Tb*CAP51 led to an effect on cell growth so that by 48 hours post RNAi induction, cell growth was reduced by half compared to the non-induced cell population. The efficiency of *Tb*CAP51 RNAi was investigated by immunoblotting, using an anti-GFP antibody to detect *Tb*CAP51::YFP expression (Fig 6.5-B). After 24 hours of *Tb*CAP51 RNAi induction, *Tb*CAP51::YFP expression was knocked down compared to non-induced controls, and by 48 hours almost no *Tb*CAP51::YFP expression was detected. Curiously, the expression of *Tb*CAP51::YFP started to recover, so that by 96 hours post RNAi induction the expression of *Tb*CAP51::YFP appeared to be equivalent in RNAi induced and non-induced cells.

To study the morphogenetic phenotype in *Tb*CAP51 RNAi induced cells, and the effect on *Tb*RP2 protein expression/localisation, cells were processed for immunofluorescence microscopy and co-labelled with the anti-*Tb*RP2 antibody (Fig 6.5-C-(1-8)). In non-induced cells, *Tb*CAP51::YFP localised to the sub-pellicular corset (as shown in Chapter 4) and cells had a normal morphology (Fig 6.5-C-(1)). However, after 24 hours of *Tb*CAP51 RNAi induction, *Tb*CAP51::YFP signal was absent from the posterior pole where new MT growth takes place during the cell division cycle. There was also accumulation of *Tb*CAP51::YFP in the middle region of the cell body, which is the posterior pole of the cell with old flagellum still positive, while the signal of the new flagellum is devoid (Fig 6.5-C-(2)). On the very same cell, it seems that kinetoplast duplication, separation has occurred, and the nucleus has undergone mitosis. However, there was no *Tb*RP2 labelling next to the new kinetoplast, which might suggest that loss of *Tb*CAP51 might has on *Tb*RP2 recruitment to the new basal body

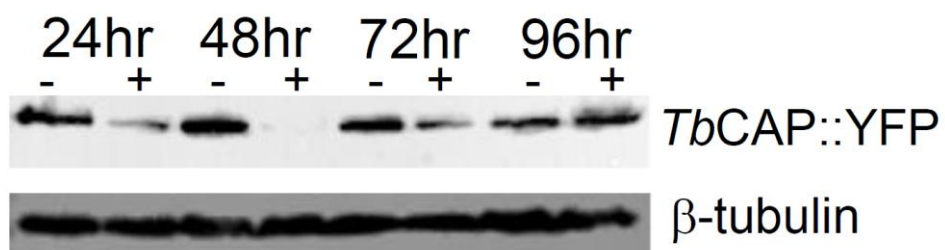
in some cells of the population (Fig 6.5-C-(2); inset). Based on the cell growth curve, even *TbCAP51* RNAi ablations introduce abnormality to the cell quickly after induction, still some cells were managed to replicate and divide. However, loss of *TbCAP51* might cause asymmetrical division, therefore, numbers of anucleate zoids cells and multi-nucleated cells within the population (Fig 6.5-C-(3)). In some of cells of the 24h induction population, due to the cytokinesis defect, cells were unable to separate. Those cells appear to be long, and the old and new flagellum at each side of the cells (Fig 6.5-C-(4-5)). After 48 hours of induction, zoid cells and cells with multi-nucleus can be observed (Fig 6.5-C-(6-8)). Kinetoplast and nuclei DAPI counts for the *TbCAP51* RNAi cell line are also analysed. Cell counts were carried out on cells from cultures, which were non-induced and 24 hours post induction of RNAi. The counts reveal that most of the cells are 1K1N in non-induced in the population and followed by 2K1N cells. After 24 h time point there is an increase number of 2K2N and zoid cells and also huge number multi-nucleated cells (Fig 6.5-D). The observation might suggest that *TbCAP51* RNAi ablation could block the mitosis and/or cytokinesis process. By 48h post induction, the majority of the cells counted were multinucleated cells followed by zoids and 1K1N cells.



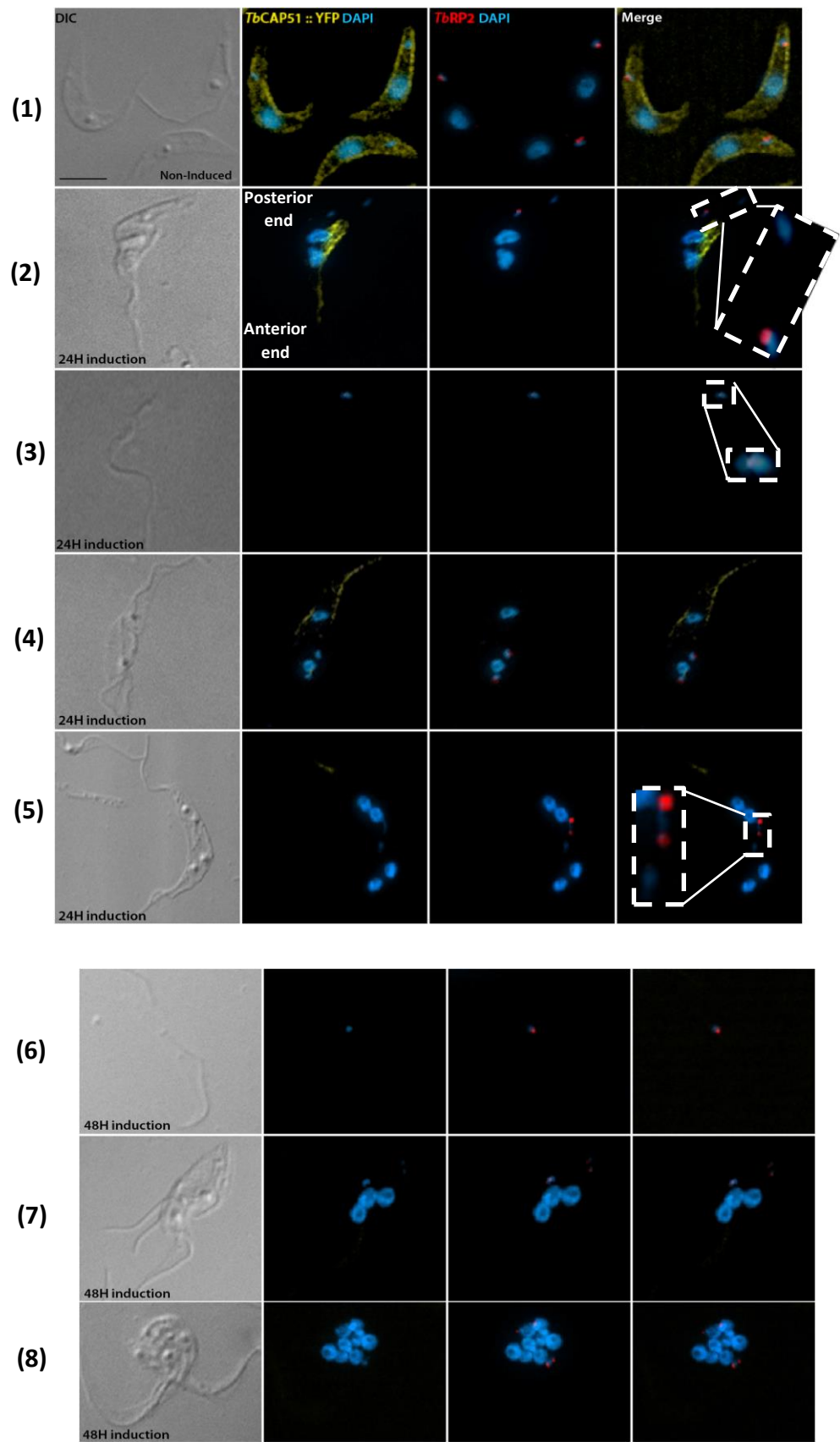
**(A)**



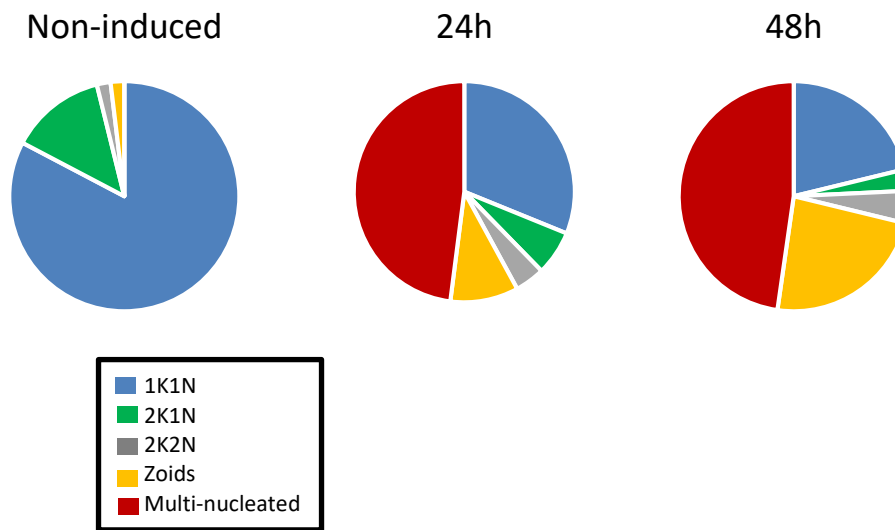
**(B)**



**(C)**



**(D)**



**Figure.6.5 RNA interference-mediated ablation of *TbCAP51* results in multiple defects in procyclic form cells. (A)** Graph showing growth of *TbCAP51*-RNAi induced (black broken line) and non-induced (dotted solid line) procyclic trypanosome cells. **(B)** The expression level of *TbCAP51*::YFP was monitored by western blotting. Both non-induced and induced cells were collected after 24h, 48, 72 and 96 h and probed with anti-GFP antibody. KMX1 detection of  $\beta$ -tubulin was included as a loading control. **(C)** Non-induced and induced cells at 24 h and 48h were collected, and stained with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m. **(D)** The *TbCAP51* RNAi cell line was induced and cells were harvested, settled onto slides and labelled with and *TbRP2* antibody. Samples were taken from a non-induced culture and cells, which were 24 and 48 hours post induction. 100 non-induced cells, 250 cells of 24h-and 100 cells of 48 hours post induction was counted based on the number of kinetoplasts and nuclei they possessed. Pie charts show an increase in multinucleated cells and zoids over time after induction.

## 6.4 T-complex protein 1 subunit epsilon (TCP-1-ε)

T-complex protein 1 subunit epsilon is part of the group II chaperonin complex called CCT (chaperonin containing TCP-1 or Tailless Complex Polypeptide 1). The TCP1 complex acts as a molecular chaperone for various proteins, including tubulin and actin, and so the identification of TCP-1-ε is intriguing in the context of its role in tubulin chaperone activity. However, TCP-1 is a large hetero-oligomeric structure, comprising multiple subunits, and so it is unclear why only the TCP 1 epsilon subunit would be biotinylated in the BioID experiments reported in this thesis.

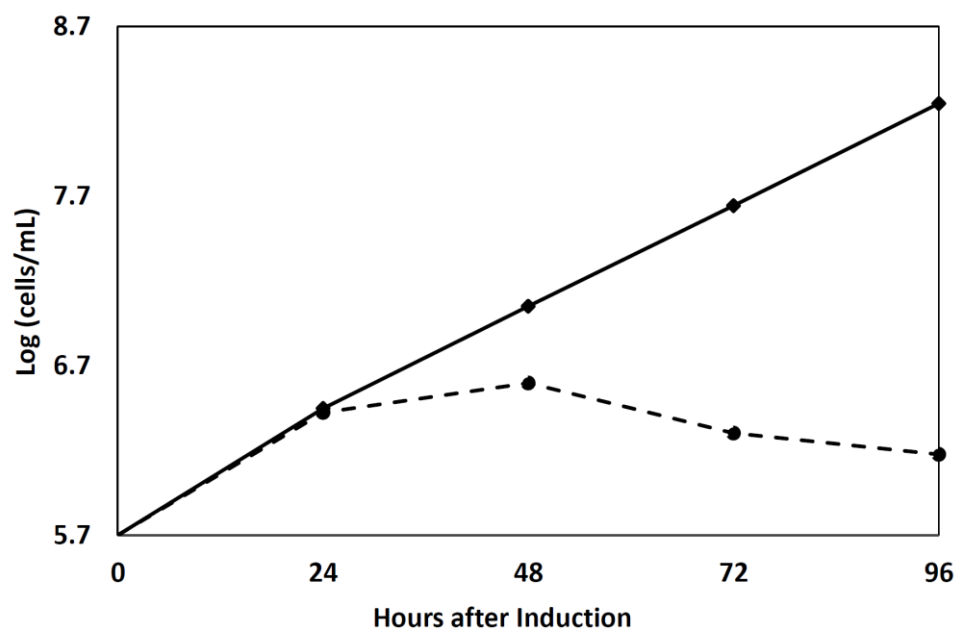
Nevertheless, to investigate the role of *Tb*TCP-1-ε, and its association with *Tb*RP2, *Tb*TCP-1-ε RNAi was induced and cell growth monitored every 24 hours (Fig 6.6-A). *Tb*TCP-1-ε RNAi induction caused a severe growth defect indicating that *Tb*TCP-1-ε is essential for cell viability. Knockdown of *Tb*TCP-1-ε expression was analysed by immunoblotting using an anti-GFP antibody to detect the *Tb*TCP-1-ε::YFP fusion protein (Fig 6.6-B). At 24 h RNAi induction, the YFP expression did not appear to be different between induced and non-induced samples. After 48 hours of *Tb*TCP-1-ε RNAi induction, expression of *Tb*TCP-1-ε::YFP was reduced in induced cells compared to non-induced controls. After 72h and 96 h after induction, the level of YFP expression came back, when compared to the induced lane at 48 hours. KMX-1 control indicating equivalence of protein loading between lanes, detection of *Tb*TCP-1-ε::YFP was variable in non-induced and RNAi-induced samples; although *Tb*TCP-1-ε::YFP expression did uniformly appear reduced in RNAi-induced cell populations compared to control samples.

To study the morphogenetic phenotypes resulting from *Tb*TCP-1-ε RNAi induction, and the effect on *Tb*RP2 expression/localisation, cells from the *Tb*TCP-1-ε RNAi depletion experiment were processed for immunofluorescence microscopy (Fig 6.6-C-(1-6)). In non-induced cells *Tb*TCP-1-ε::YFP expression was evident throughout the cell body as expected and cells showed normal morphology (Fig 5.6-C-(1)). However, after 24 hours of *Tb*TCP-1-ε RNAi induction, the *Tb*TCP-1-ε::YFP cell body localisation was

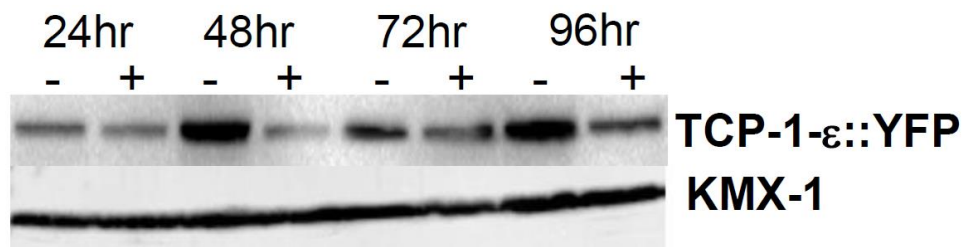
severely affected, the YFP signal was no longer uniformly dispersed throughout the cytoplasm, and the *TbTCP-1-ε::YFP* signal was focussed into distinct foci (Fig 6.6-C-(2-5)). In cells in which expression of *TbTCP-1-ε* was targeted by RNAi induction, kinetoplasts and nuclei were clearly still able to undergo replication and segregation, however their overall morphology was abnormal. The mid-point of many cells was expanded (Fig 6.6-C-(2); black arrow) as early as 24 hours post RNAi induction. Cells which showed a 'monstrous' appearance also accumulated, i.e. cells with multiple nuclei and abnormal kinetoplast complement indicating defects in cell division developed in the *TbTCP-1-ε* depleted cells. *TbTCP-1-ε::YFP* RNAi also produced cells with detached flagella and kinetoplasts that were incorrectly positioned with respect to nuclei (Fig 6.6-C-(3-4); white arrows). The cytokinesis defects were clearly evident in *TbTCP-1-ε* depleted cells which showed abnormally long posterior ends or cells which remained attached via their posterior ends (Fig 6.6 C-(5-6)). Although loss of *TbTCP-1-ε* protein cells resulted in severe morphogenetic defects the expression and localisation of *TbRP2* remained unaffected. However, in a small number of cells *TbRP2* was detected in a region of the cell not associated with kinetoplast DNA (Fig. 6.6 C-(3-4), white asterisks), suggesting that *TbTCP-1-ε* depletion could disturb the connection between the mature basal body (detected by *TbRP2*) and kinetoplast DNA.

In the previous experiments, it was difficult to distinguish that a flagellum was present on some cells. To investigate whether by *TbTCP-1-ε* knock-down affected flagellum assembly, cells were analysed by immunostaining cells with the anti-PFR2 monoclonal antibody L8C4 (Fig. 6.7 D-(1-5)). These experiments demonstrated that the majority of the cells induced for *TbTCP-1-ε* RNAi were able to assemble flagella but confirmed the flagella detachment phenotype. This might suggest that loss of *TbTCP-1-ε* disrupts perturbs flagellum-cell body attachment.

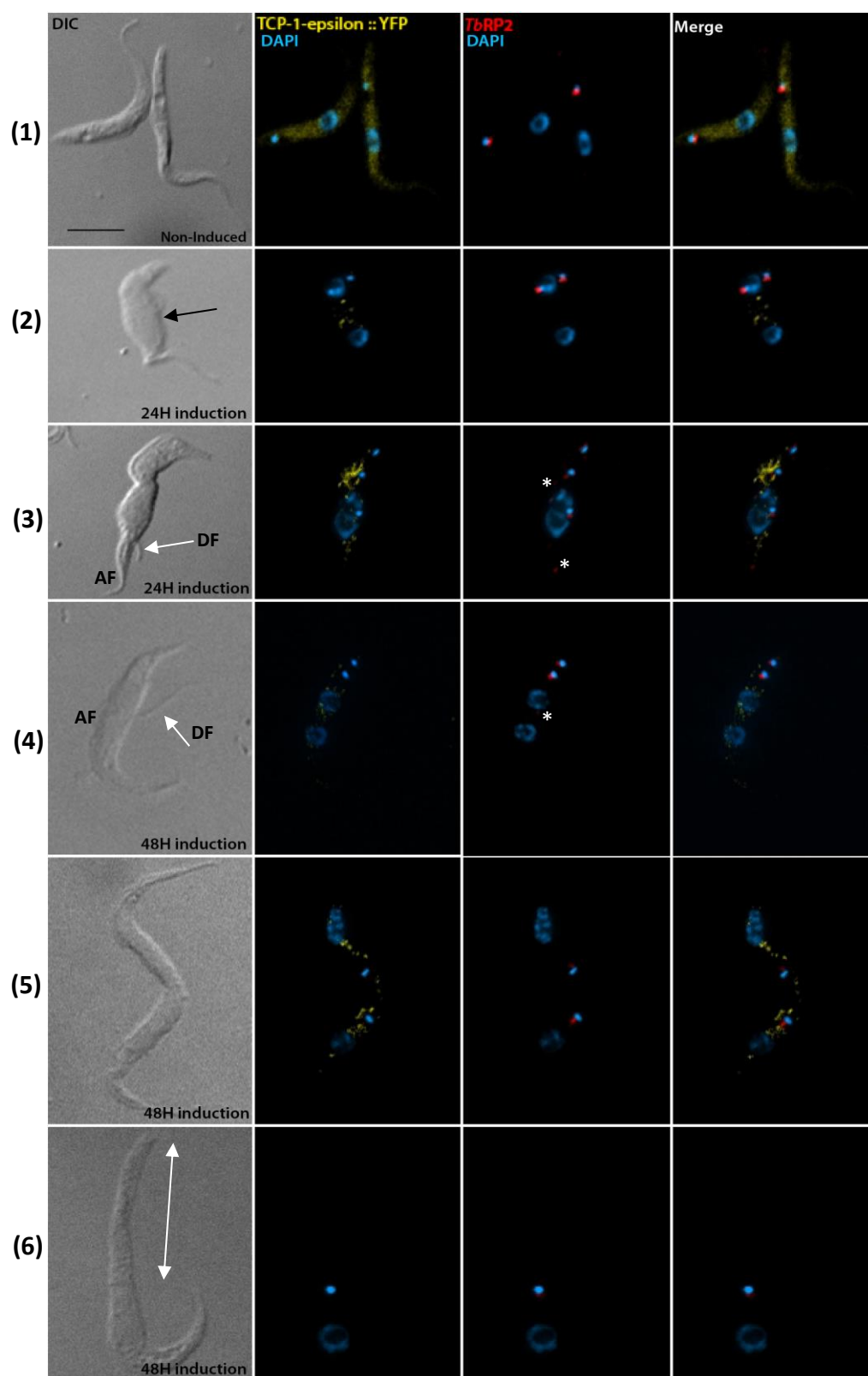
**(A)**



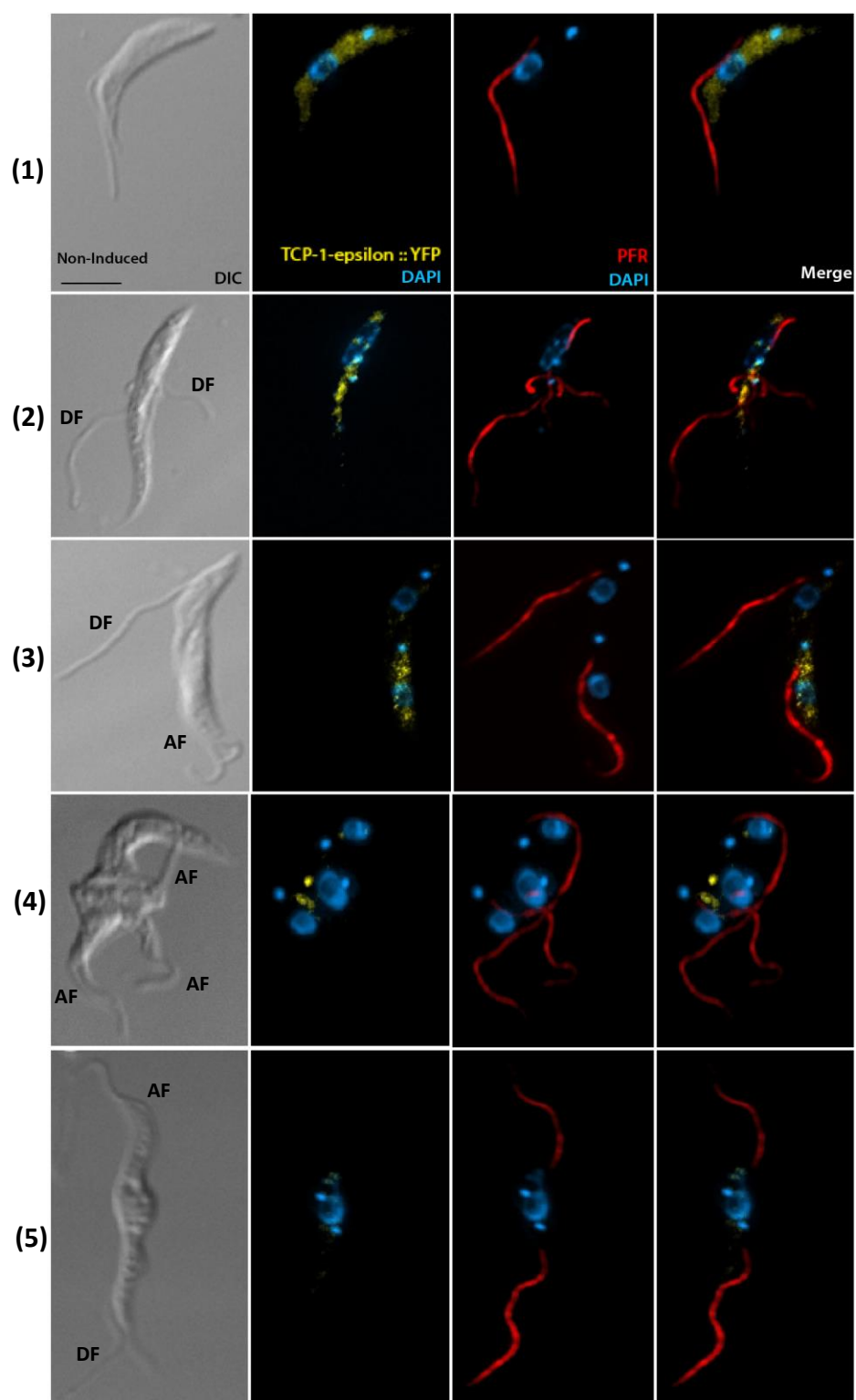
**(B)**



**(C)**



**(D)**





**Figure 6.6 RNA interference-mediated ablation of *TbTCP-1-ε* results in multiple defects in procyclic form cells. (A)** Graph showing growth of *TbTCP-1-ε* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. **(B)** The *TbTCP-1-ε* expression level was monitored by western blotting. Both non-induced and induced cells were collected 24h, 48 h, 72h and 96h post induction and probed with anti-GFP antibody. KMX1 detection of  $\beta$ -tubulin was included as a loading control. **(C)** Non-induced and induced cells at 24h/48h post induction of *TbTCP-1-ε* RNAi cell line were harvested for immunofluorescence analysis. Cells were stained with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. **(D)** *TbTCP-1-ε* RNAi cell line were harvested for immunofluorescence analysis. Cells were stained with L8C4 to label the flagella. Scale bar = 10  $\mu$ m. **AF** attached flagellum; **DF** detached flagellum.

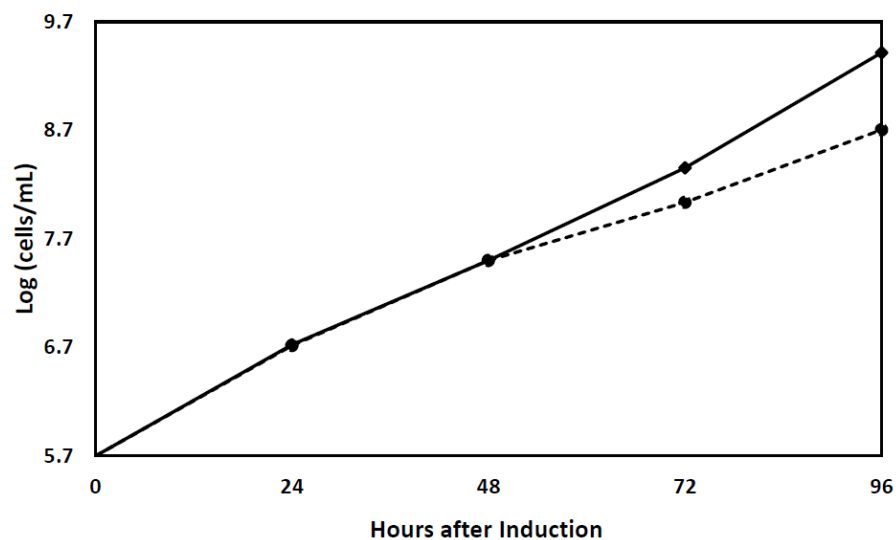
## 6.5 ADP-ribosylation factor-like 3 (ARL3)-like

The identification of an ARL3-like protein as a putative *TbRP2*-interacting/near neighbour protein was anticipated given that; (i) human RP2 protein is reported to act as a GAP for an ARL3 protein and, (ii) it has previously been reported that RNAi knockdown of an ARL3-like protein (*TbARL3A*) in *T. brucei* leads to a flagellum assembly defect (Sahin et al., 2004), similar to that observed following RNAi ablation of *TbRP2*. However, the ARL3-like protein studied by Sahin et al (Sahin et al., 2004) and the ARL3-like protein identified in the BioID study reported here are not the same protein. Thus, to investigate the role(s) of the *TbARL3*-like protein identified in this thesis, *TbArl3L* RNAi was induced in procyclic cells and cell growth monitored every 24 hours (Fig 6.7-A). Little difference in growth was observed for the first 48 hours following RNAi induction, but thereafter cell growth was impaired, until the termination of the induction experiment at 96 hours. Knockdown of *TbArl3L* expression was confirmed by immunoblotting using the anti-GFP antibody to detect *TbArl3L::YFP* with detection of  $\beta$ -tubulin by KMX-1 used as a loading control (Fig 6.7-B). After 24 hours of RNAi induction, *TbARL3L::YFP* expression was reduced in RNAi induced cells, and after 48 hours very little *TbARL3L::YFP* expression could be detected.

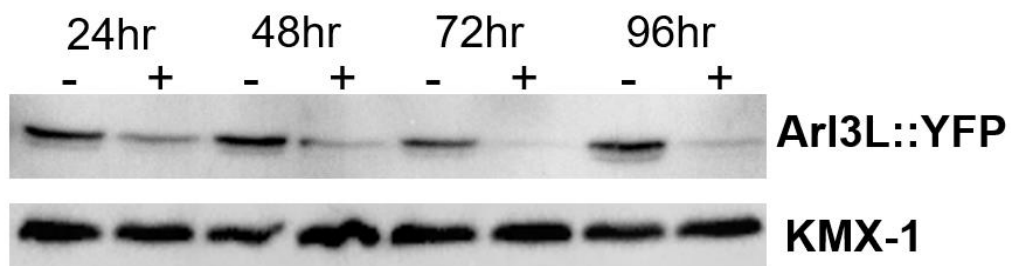
To study the morphogenetic phenotypes resulting from *TbARL3L* RNAi ablation, cells from a non-induced culture and from 24, 48, 72 and 96 hours post RNAi induction were processed for immunofluorescence microscopy. The cells were co-labelled with anti-*TbRP2* antibody to investigate the effect *TbARL3L* RNAi had on the expression and basal body localisation of *TbRP2* (Fig 6.7-C-(1-9)). In Fig 6.7-C-1, non-induced cells clearly show *TbARL3::YFP* co-localised at the basal body with *TbRP2*, but also to the nucleus where the *TbARL3::YFP* signal was diffuse throughout the nucleus. After 24 hours of RNAi induction, in most cells *TbARL3::YFP* signal was undetectable at the basal body and nucleus, although the morphology of these cells appeared normal (Fig 6.7-C-2). However, by 48 hours post induction, a number of cells displayed a detached flagellum (Fig 6.7-C-(3-4); white arrows), and the flagellar pocket may be enlarged in

approximately 20% of cells (Fig 6.7-C-3, 4 and 6; white square). One of the consequences of having a detached flagellum is that the *T. brucei* cells are unable to divide correctly and organelle positioning is affected. Thus, multinucleate cells and kinetoplast/nuclei mispositioning defects were prevalent in the population (Fig. 6.7-C-5, 7, 8 and 9). In addition, depletion of *TbARL3L* also led to cytokinesis defects, with cells apparently unable to divide and remaining attached via their posterior pole (Fig. 6.7-C-6; white asterisk). Despite these gross morphological defects, in the majority of cells *TbRP2* was expressed and localised adjacent to the kinetoplast, consistent with a mature basal body localisation. However, closer scrutiny revealed that *TbRP2* labelling was (i) not associated with all kinetoplasts (Fig.6.7-C-5) and (ii) present at locations within the cell where kinetoplasts were absent. This suggests the linkage between the basal body and kinetoplast DNA, established by the TAC, was also affected in some cells following ablation of *TbARL3L*.

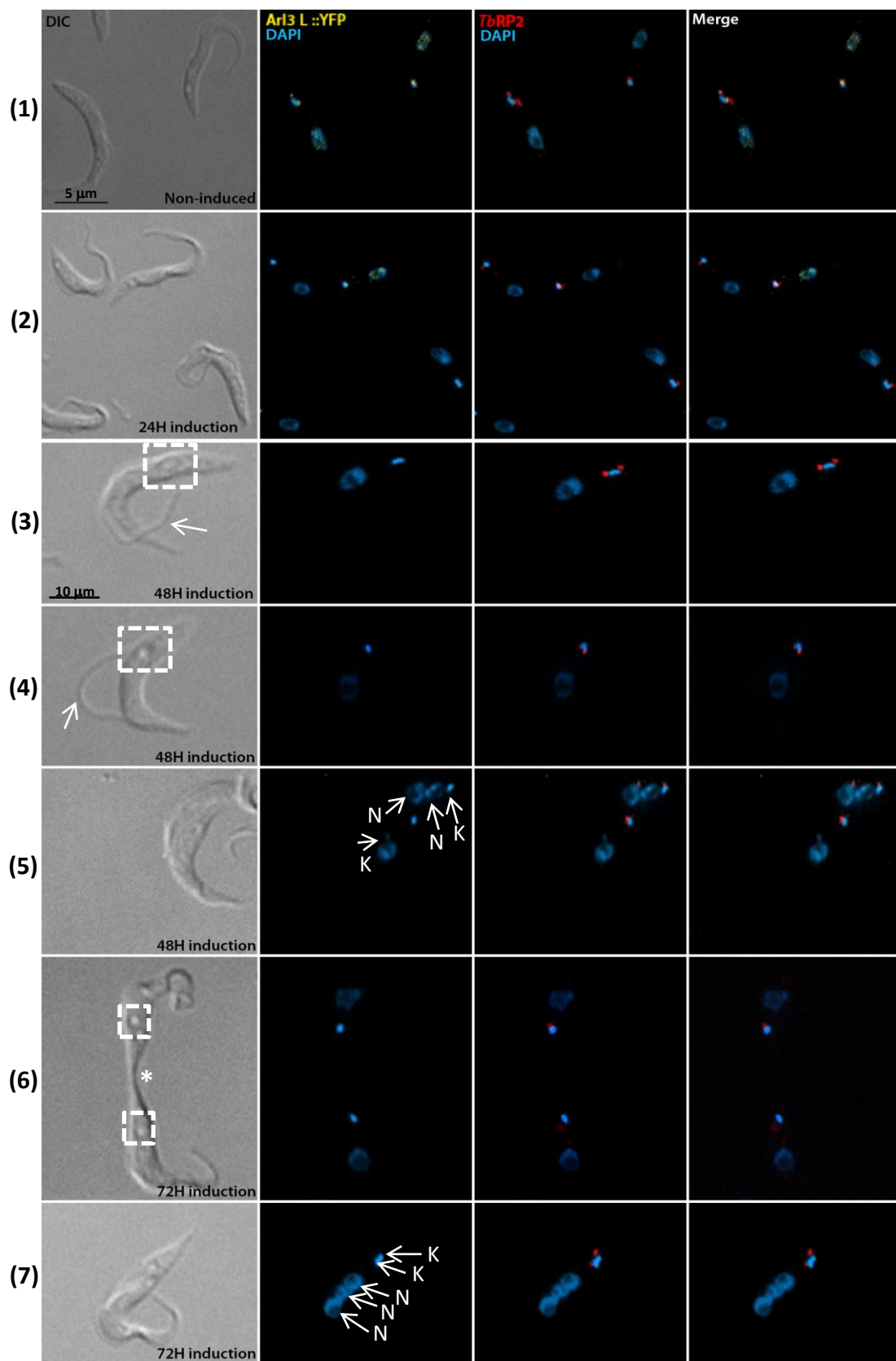
**(A)**

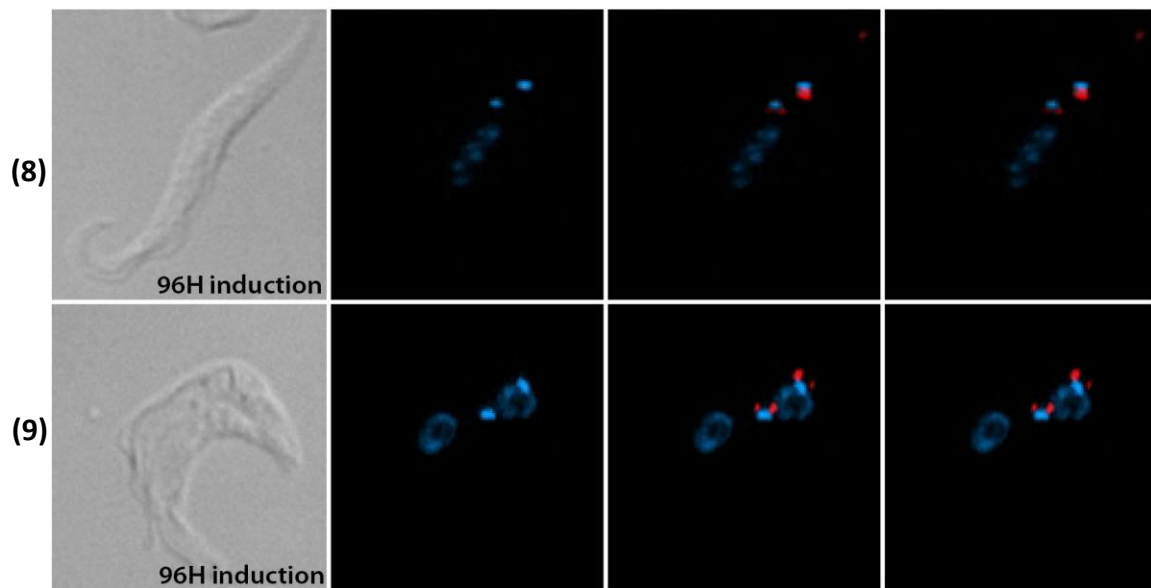


**(B)**



**(C)**



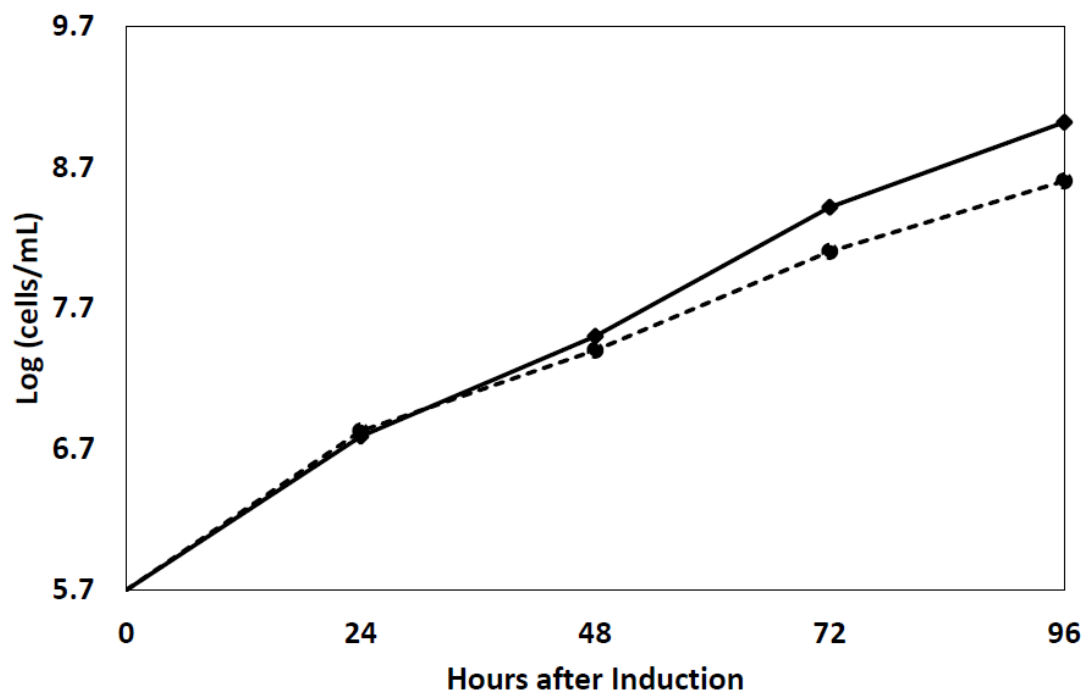


**Figure 6.7 The effect of RNA interference-mediated ablation of *TbARL3L* on procyclic form cells.** (A) Graph showing growth of *TbARL3L* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. (B) The *TbARL3L* expression level was monitored by western blotting. Both non-induced and induced cells were collected 24h, 48 h, 72h and 96h post induction and probed with anti-GFP antibody. KMX1 detection of  $\beta$ -tubulin was included as a loading control. (C) Non-induced and induced cells at 24h, 48 h, 72h and 96h post induction of *TbTCP-1- $\epsilon$*  RNAi cell line were harvested for immunofluorescence analysis. Cells were stained with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. White arrow indicates flagellum detachment, white square indicates enlarged flagellar pocket, and white asterisk cell unable to divide and remaining attached via their posterior pole.

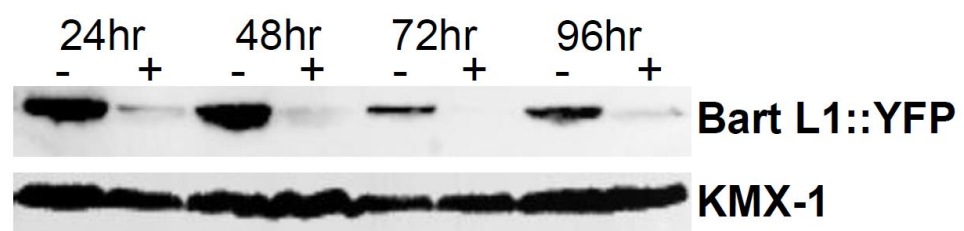
## 6.6 ARL2-binding protein (BART, Binder of ARL Two) like-1

The identification of a BART-like1 protein was particularly intriguing given that during the course of my PhD studies, it was published that BARTL1 has a regulatory role in ciliary assembly; potentially aiding the presentation of active ARL3 to RP2 (or hindering ARL3 membrane binding at the transition zone). To investigate the role(s) of *TbBARTL1*, its expression was knocked down by RNAi and cell growth monitored every 24 hours (Fig 6.8-A). This revealed only a modest effect on cell growth. The expression of *TbBARTL1::YFP* was monitored by immunoblotting using the anti-GFP antibody, with detection of  $\beta$ -tubulin by KMX-1 acting as a loading control (Fig 6.8-B). This experiment confirmed *TbBARTL1::YFP* expression was almost undetectable within 24 hours of initiating RNAi induction. Fluorescence microscopy also confirmed the efficacy of *TbBARTL1* RNAi knockdown, with *TbBARTL1::YFP* significantly reduced within 24 hours. Despite the robust knockdown of *TbBARTL1* expression, no gross phenotypic abnormalities were observed in RNAi induced cells; even after prolonged (96 hours) of RNAi induction (Fig 6.8-C-(1-5)). A small proportion of cells showed some abnormality e.g. detached flagella and multinucleate phenotype (Fig 6.8-D-(1-3)) which might explain the modest effect on cell growth, but as all cells showed *TbBARTL1* depletion it is uncertain why phenotypic abnormalities would only be present in a small proportion of the cells. What was clear from these experiments was that loss of *TbBARTL1* expression had no effect on expression or basal body localisation of *TbRP2*.

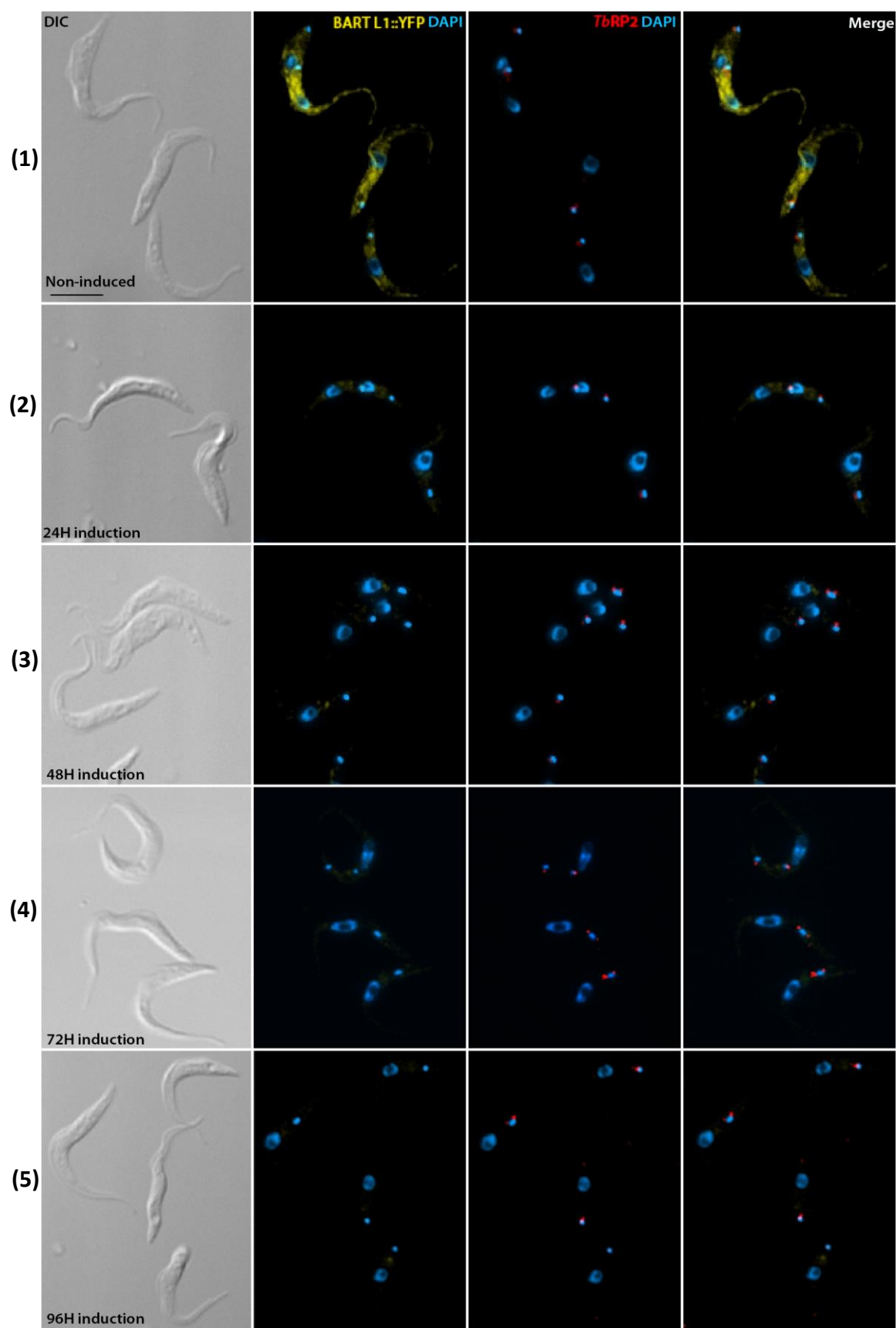
**(A)**



**(B)**

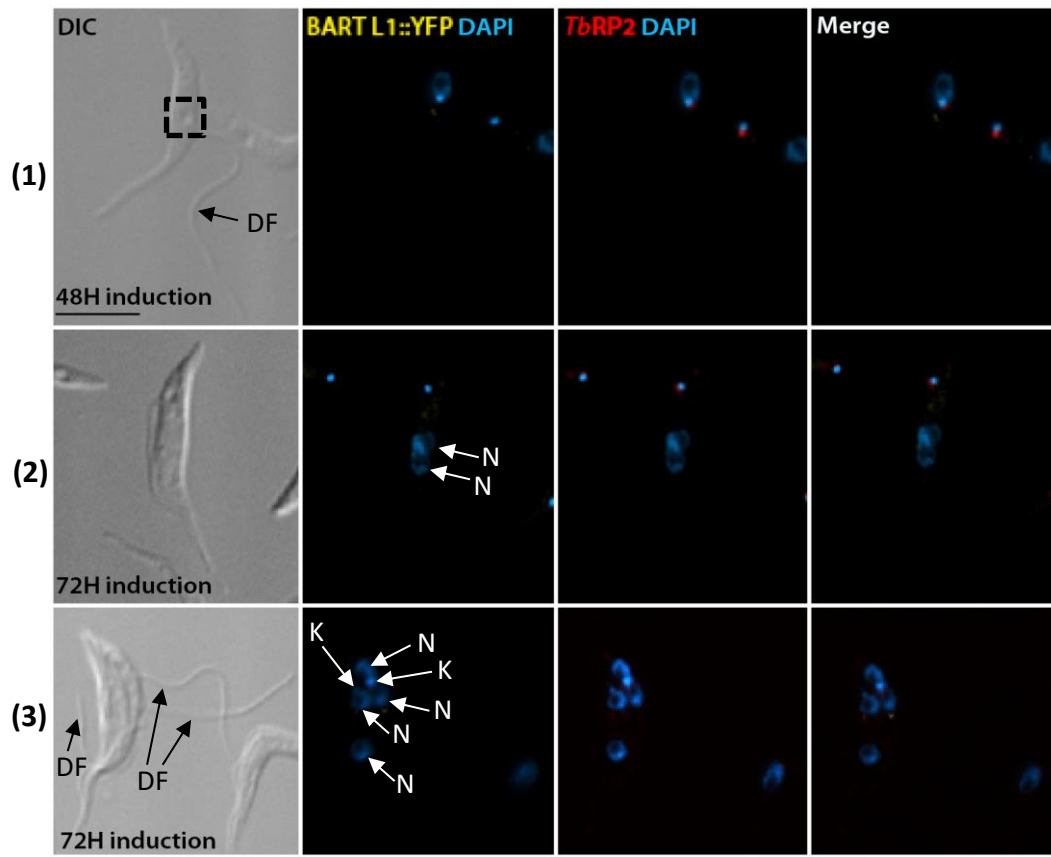


**(C)**





**(D)**



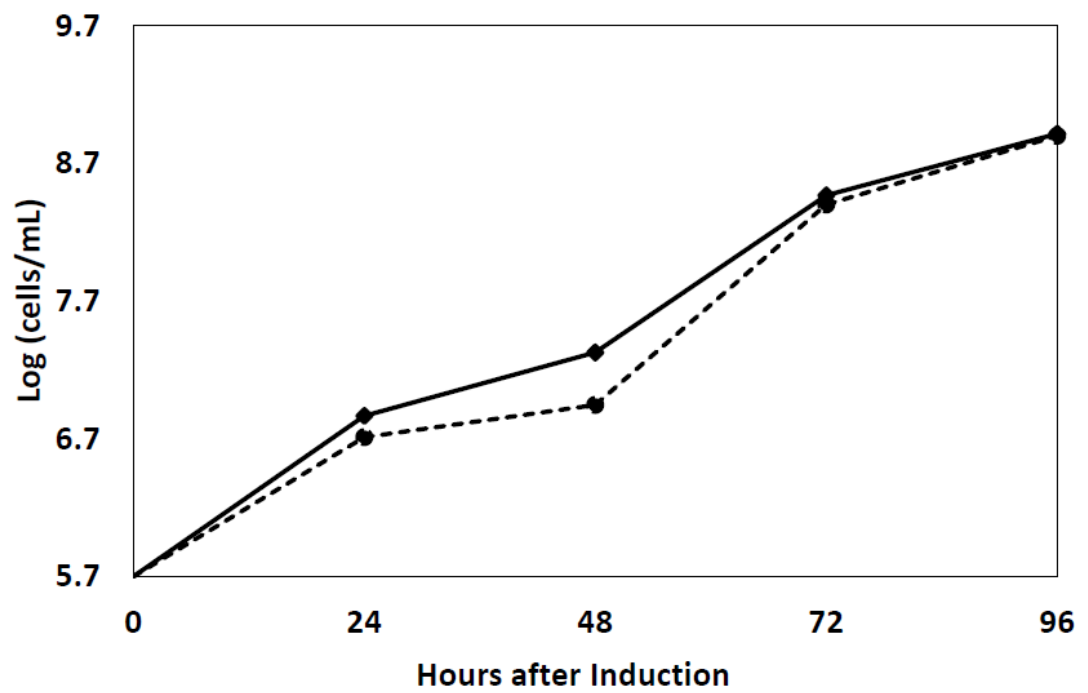
**Figure 6.8. The effect of RNA interference-mediated ablation of *TbBARTL1* on procyclic form cells.** (A) Graph showing growth of *TbBARTL1* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. (B) The *TbBARTL1* expression level was monitored by western blotting. Both non-induced and induced cells were collected 24h, 48 h, 72h and 96h post induction and probed with anti-GFP antibody. KMX1 detection of  $\beta$ -tubulin was included as a loading control. (C) Non-induced and induced cells at 24h, 48 h, 72h and 96h post induction of *TbBARTL1* RNAi cells were harvested for immunofluorescence analysis. (D) Immunofluorescence analysis revealed that small number of cells were appeared with cell division defect and flagellum detachment after *TbBARTL1* RNAi ablation. Cells were stained with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m. DF-detached flagellum.

## 6.7 Heat Shock Protein 70 (HSP70)

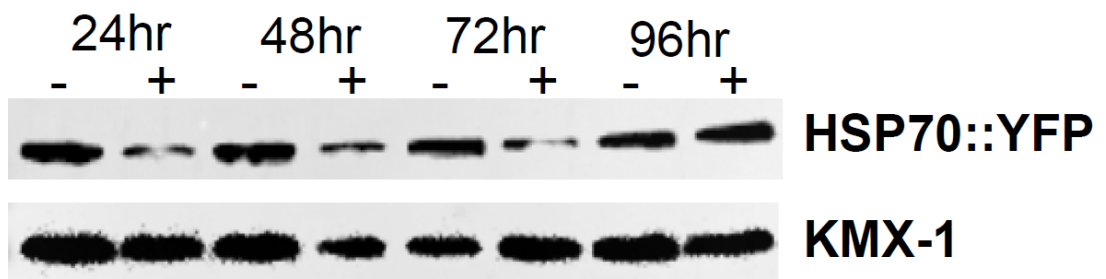
In this chapter, I focus on the effect(s) that *TbHSP70* RNAi ablation has on *TbRP2* protein expression and localisation. The candidate protein *TbHSP70* is a detergent-insoluble protein, and *TbHSP70::YFP* localised to the basal body, which co-localised at with *TbRP2* at 1K1N stage. There is also an intense focus of *TbHSP70::YFP* signal localised distal to *TbRP2*, and this YFP signal continues along on the flagellum. As the cell cycle progresses, the *TbHSP70* signal at basal body and old flagellum diminishes, however, the *TbHSP70* signal distal to the basal body duplicates and segregates with the replicated kinetoplast.

To investigate the putative association between *TbHSP70* and *TbRP2* in *T. brucei* (as indicated by the BioID experiments), the expression of *TbHSP70* was ablated by RNAi and cell growth monitored every 24 hours in RNAi induced and non-induced cell populations (Fig 6.9-A). Although cell growth appeared to be impaired within the first 48 hours of RNAi induction, cell growth unexpectedly recovered from 48 hours post induction. The expression of *TbHSP70::YFP* was monitored by immunoblotting using the anti-GFP antibody to detect *TbHSP70::YFP*, with detection of  $\beta$ -tubulin by KMX-1 used as a loading control (Fig 6.9-B). The immunoblotting experiments revealed that expression of *TbHSP70::YFP* was successfully knocked down by RNAi up to 72 hours post induction, but that at 96 hours, expression of *TbHSP70::YFP* had recovered to wild type levels. Immunofluorescence analysis confirmed that expression of *TbHSP70::YFP* was knocked down within the 72 hours of RNAi ablation. At 96 hours post induction, a distinct focus of *TbHSP70::YFP* distal to the basal body was detected by fluorescence microscopy. Despite the apparent efficacy of *TbHSP70* RNAi knockdown (at least up to 72 hours post induction), no obvious phenotype could be observed (Fig 6.9-C-(1-5)), nor did the loss of *TbHSP70* have any effect on the expression or localisation of *TbRP2* (with *TbRP2* clearly detectable in a position close to the kinetoplast in all cells).

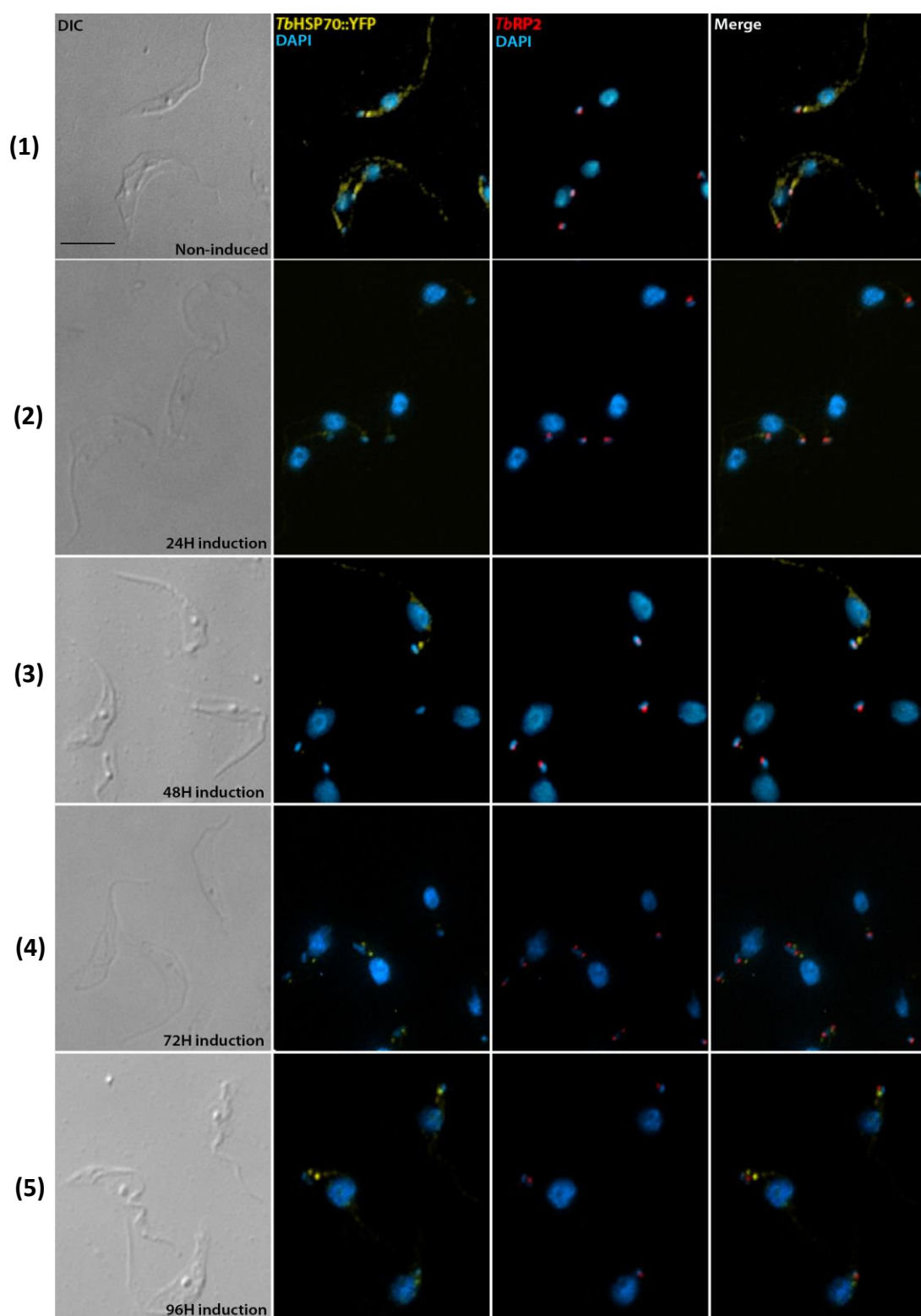
**(A)**



**(B)**



**(C)**



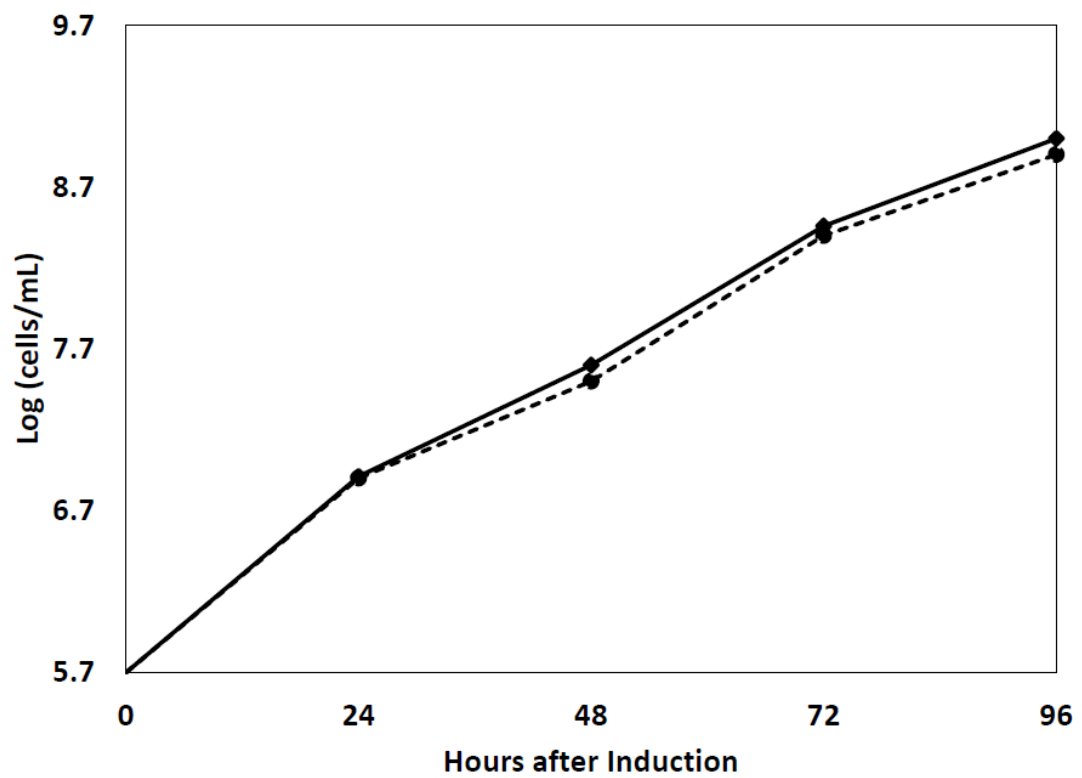
**Figure 6.9. The effect of RNA interference-mediated ablation of *TbHSP70* on procyclic form cells.** (A) Graph showing growth of *TbHSP70* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. (B) The *TbHSP70* expression level was monitored by western blotting. Both non-induced and induced cells were collected at 24h, 48 h, 72h and 96h post induction and probed with anti-GFP antibody. KMX1 detection of  $\beta$ -tubulin was included as a loading control. (C) Non-induced and induced cells at 24h, 48 h, 72h and 96h post induction of *TbHSP70* RNAi cells were harvested for immunofluorescence analysis. Cells were stained with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m.

## 6.8 Vps-27, Hrs and STAM (VHS) domain containing protein

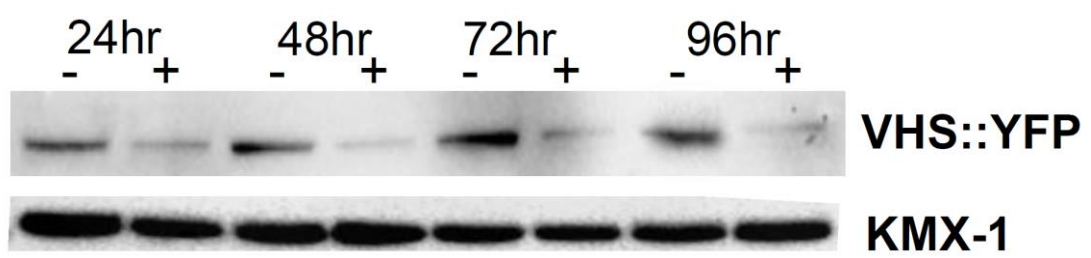
The identification of a VHS domain-containing protein as a putative TbRP2-interacting/near neighbour protein was intriguing, given the proposed role(s) of VHS domain-containing proteins in vesicular trafficking in other eukaryotes. To investigate the role(s) of the *TbVHS* domain-containing protein, expression of *TbVHS* was ablated by RNAi and cell growth monitored every 24 hours (Fig 6.10-A). No growth defect was detected following RNAi knockdown of the *TbVHS* protein. The knockdown of *TbVHS::YFP* expression was monitored by immunoblotting, using the anti-GFP antibody, with detection of  $\beta$ -tubulin by KMX-1 used as a loading control (Fig 6.10-B). The immunoblotting experiment indicated that RNAi mediated ablation of *TbVHS::YFP* was successful.

RNAi ablation of *TbVHS* was also investigated by fluorescence microscopy (Fig 6.10-C-(1-5)). In non-induced cells (Fig 6.10-C-1), *TbVHS::YFP* can be observed accumulating in the region between the kinetoplast and nucleus, and the *TbVHS::YFP* signal that appear to be more distal to *TbRP2*. After induction of *TbVHS* RNAi, the *TbVHS::YFP* signal is progressively diminished, but no obvious phenotype can be observed, even 96 hours after induction of RNAi (Fig 6.10-C-(2-5)). Moreover, depletion of *TbVHS* from the cell has no effect on *TbRP2* expression or localisation.

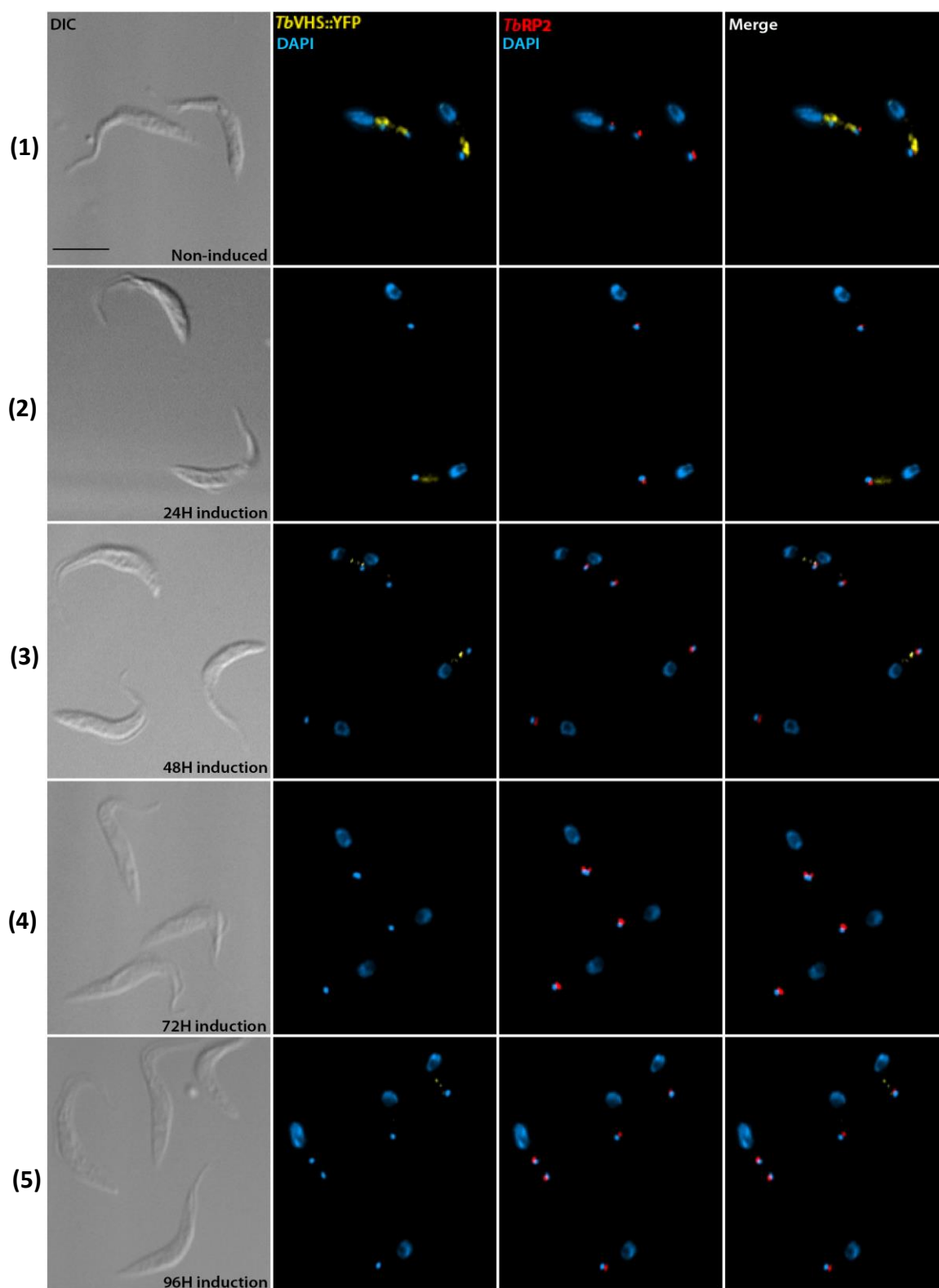
**(A)**



**(B)**



**(C)**





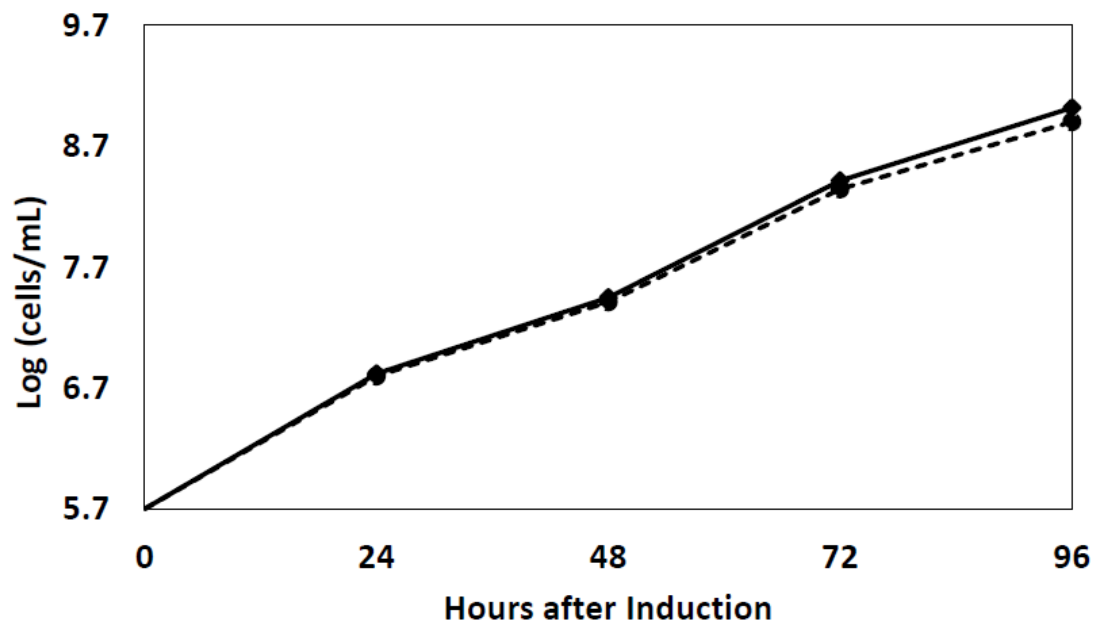
**Figure 6.10 The effect of RNA interference-mediated ablation of *TbVHS* on procyclic form cells.** (A) Graph showing growth of *TbVHS* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. (B) The *TbVHS* expression level was monitored by western blotting. Both non-induced and induced cells were collected at 24h, 48 h, 72h and 96h post induction and probed with anti-GFP antibody. KMX1 detection of  $\beta$ -tubulin was included as a loading control. (C) Non-induced and induced cells at 24h, 48 h, 72h and 96h post induction of *TbVHS* RNAi were harvested for immunofluorescence analysis. Cells were stained with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m.

## 6.9 Protein kinase

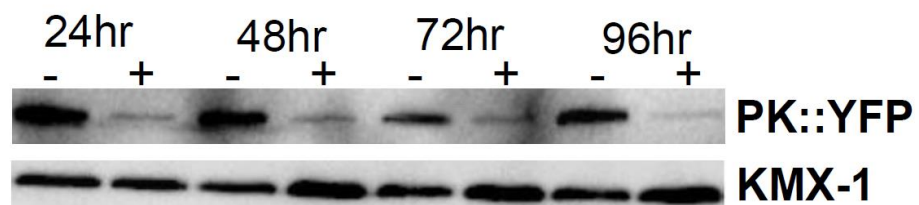
To investigate the putative function(s) of the protein kinase (*TbPK*) identified in the BiolD experiments its expression was knocked by gene specific RNAi and growth of RNAi induced and non-induced cells monitored every 24 hours (Fig 6.11-A). Based on cell growth data, the RNAi knockdown of this protein kinase has no effect on the growth and proliferation of procyclic form cells. RNAi knockdown of *TbPK* expression was confirmed by immunoblotting using anti-GFP antibody to detect *TbPK::YFP*; with KMX-1 detection of  $\beta$ -tubulin used as a loading control (Fig 6.11-B). After 24 hours of RNAi induction, *TbPK::YFP* expression was reduced in RNAi induced cells, and after 96 hours very little *TbPK::YFP* expression could be detected in the RNAi induced population.

*TbPK* RNAi induced (and non-induced) cells were also analysed by fluorescence microscopy to detect *TbPK::YFP* expression and *TbRP2* (Fig 6.11-C-(1-5)). In non-induced cells (Fig 6.11-C-1), *TbPK::YFP* can be seen localised close to the kinetoplast; aligning in part with *TbRP2* at the basal body but also extending to a position more distal to *TbRP2*. However, although 24 hours after initiation of *TbPK* RNAi induction, *TbPK::YFP* expression was successfully knocked down no obvious phenotype was observed, even after extended (96 hour) induction (Fig 6.11-C-5). Loss of *TbPK* also has no effect on *TbRP2* expression or localisation.

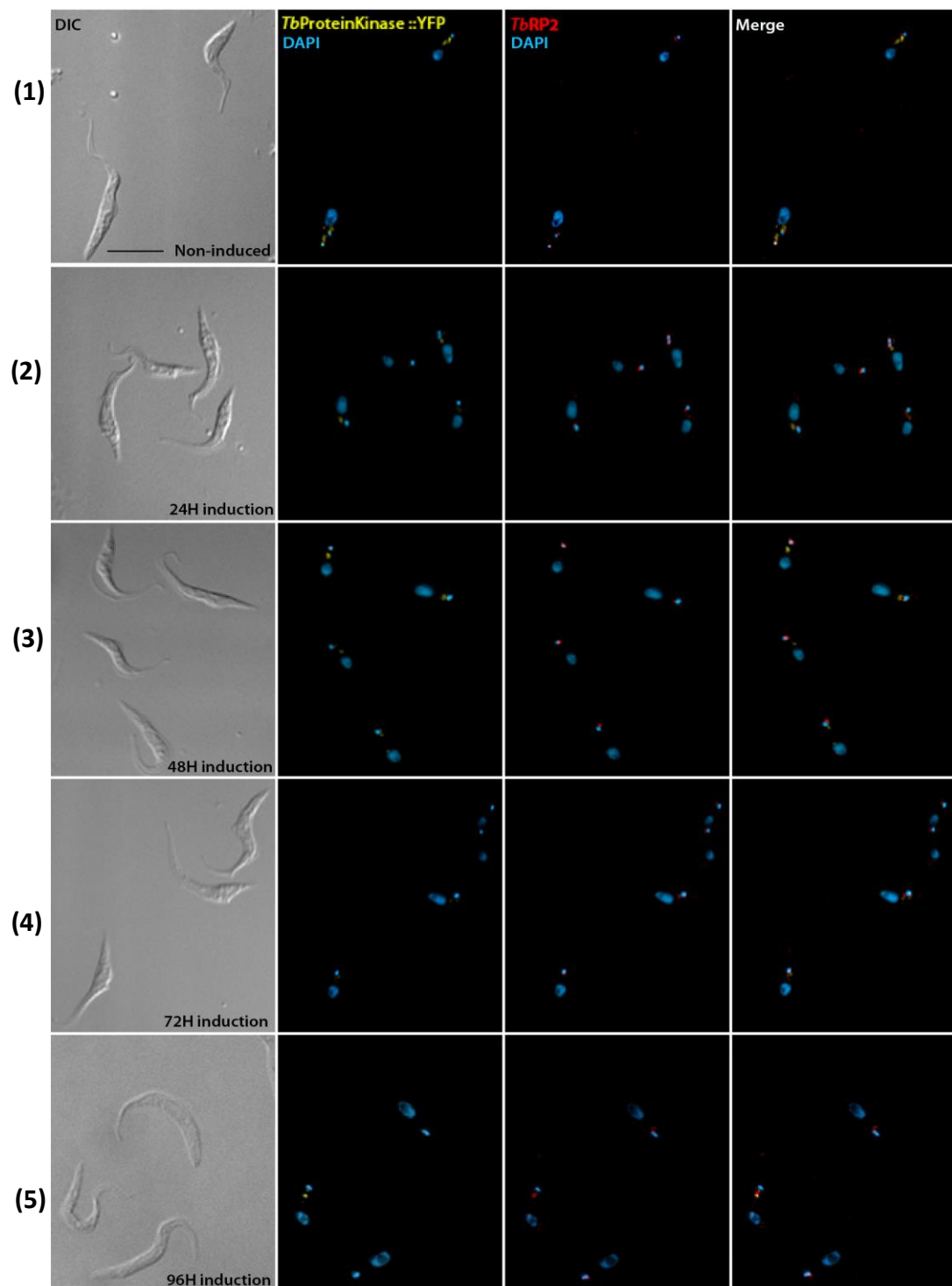
**(A)**



**(B)**



**(C)**

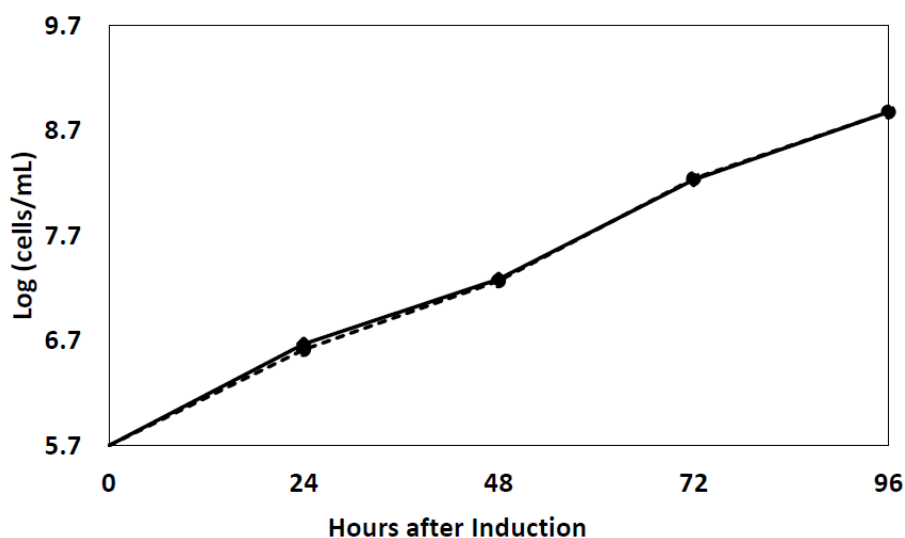


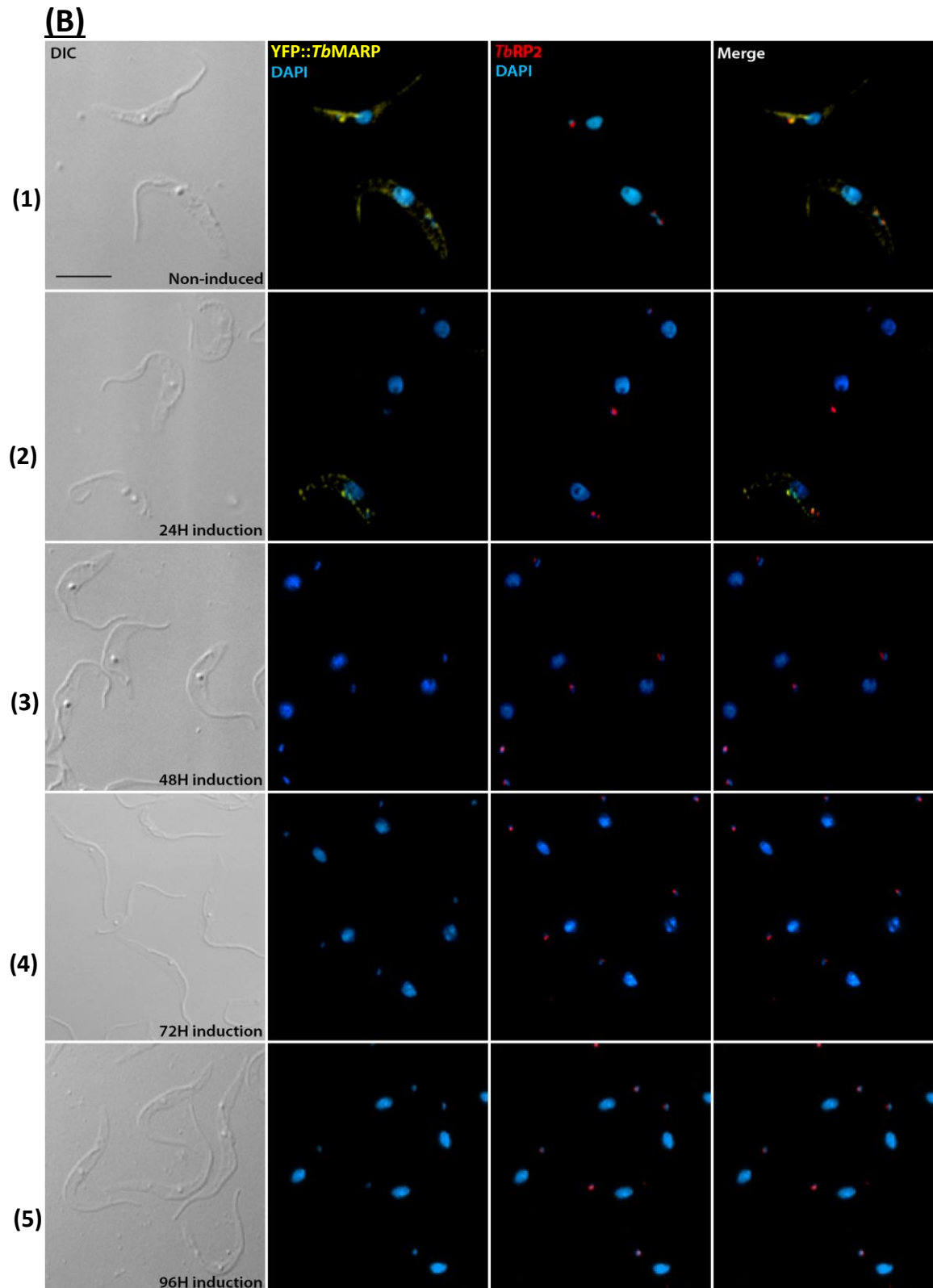
**Figure 6.11 The effect of RNA interference-mediated ablation of *TbPK* on procyclic form cells.** (A) Graph showing growth of *TbPK* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. (B) The *TbPK* expression level was monitored by western blotting. Both non-induced and induced cells were collected at 24h, 48 h, 72h and 96h post induction and probed with anti-GFP antibody. KMX1 detection of  $\beta$ -tubulin was included as a loading control. (C) Non-induced and induced cells at 24h, 48 h, 72h and 96h post induction were harvested for immunofluorescence analysis. Cells were stained with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m.

## 6.10 Microtubule-associated repetitive protein (MARP)

To investigate the putative function(s) of the microtubule associated repetitive protein (MARP) identified by *Tb*RP2 BioID, the protein was depleted by RNAi and the effect on cell growth monitored every 24 hours (Fig 6.12-A). Based on cell counts, *Tb*MARP RNAi knockdown has no effect on the proliferation and growth of procyclic form cells. Detection of YFP::*Tb*MARP expression by western blotting was not possible due to its very large molecular weight (attempts to resolve this protein on 8% SDS-PAGE were unsuccessful). However, expression of YFP::*Tb*MARP was analysed by fluorescence microscopy (Fig 6.12-B-(1-5)). Non-induced cells are shown on the top panel (Fig 6.12-B-1), and YFP::*Tb*MARP is localised (as expected) at the basal body (co-localised with *Tb*RP2), but also with the sub-pellicular microtubule cytoskeleton. After 24 hours of RNAi induction, YFP::*Tb*MARP signal is depleted in most cells. At later time points (48-96 hours), YFP::*Tb*MARP was unable to be detected in RNAi induced cells, however, no gross morphological phenotype could be observed (even after 96 hours of RNAi induction). However, RNAi cells collected at 72 and 96 hours post induction looked more elongate and thinner than non-induced cells and to RNAi induced cells collected at early RNAi time points (Fig 6.12-B-(2-5)). However, loss of *Tb*MARP from the cell has no effect on either *Tb*RP2 expression or basal body localisation.

**(A)**



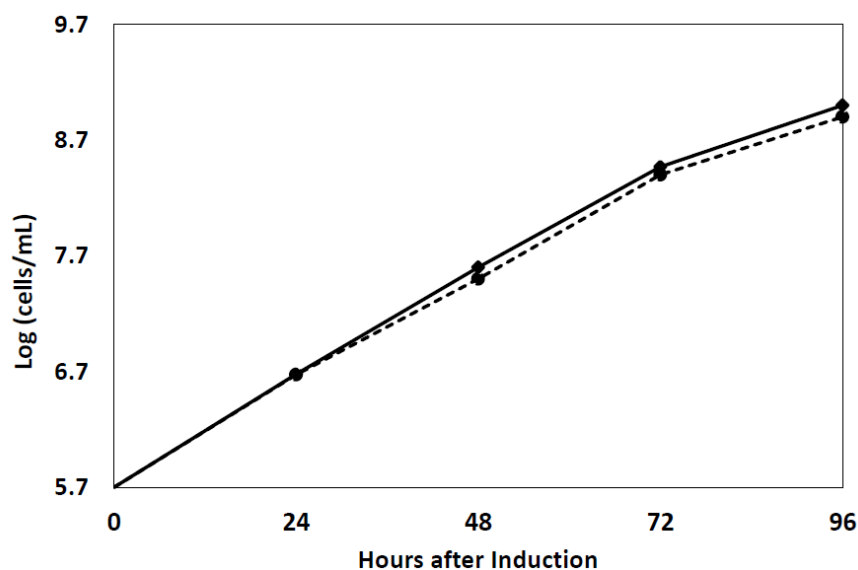


**Figure 6.12 The effect of RNA interference-mediated ablation of *TbMARP* on procyclic form cells.** (A) Graph showing growth of *TbMARP* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. (B) Non-induced and induced cells at 24h, 48 h, 72h and 96h post induction were harvested for immunofluorescence analysis. Cells were stained with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m.

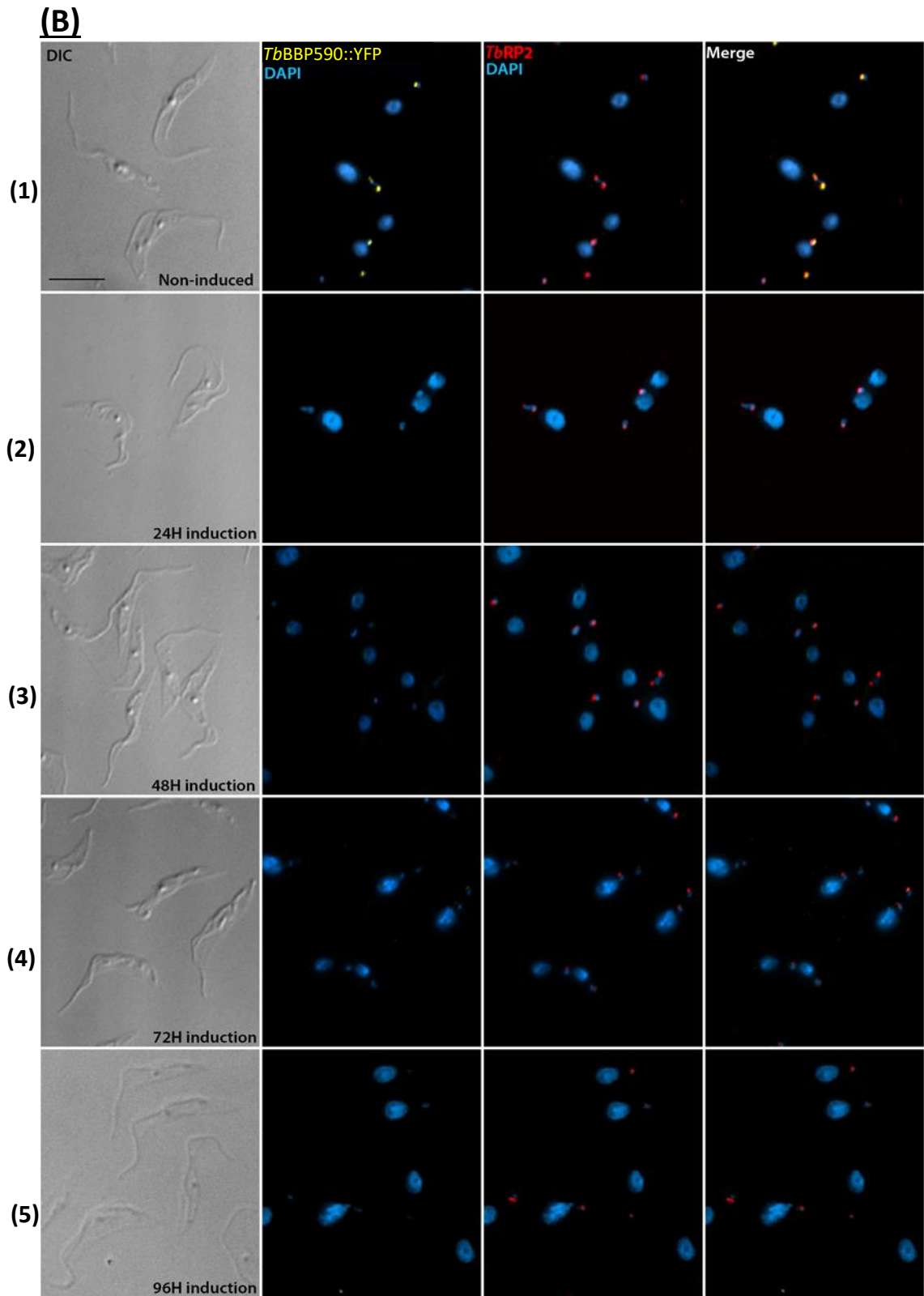
## 6.11 *TbBB590*

To investigate the putative function(s) of *TbBB590*, the protein was depleted by RNAi and the effect on cell growth monitored every 24 hours (Fig 6.13-A). Based on cell counts, *TbBB590* RNAi knockdown has no effect on the proliferation and growth of procyclic form cells (there is only slightly cell number drop between 72h to 96h of induction). Detection of *TbBB590::YFP* expression by western blotting was not possible due to its large molecular weight (all attempts to resolve this protein by SDS-PAGE were unsuccessful). However, *TbBB590::YFP* expression could be analysed by fluorescence microscopy in RNAi induced and non-induced cells (Fig 6.13-B-(1-5)). In non-induced cells (Fig 6.13-B-1), *TbBB590::YFP* is localised at the basal body, co-localised with *TbRP2*. As previously described (Chapter 4), *TbBB590::YFP* localisation at the mature basal body has an unusual (unprecedented?) cell cycle pattern; i.e. *TbBB590::YFP* expression is more intense on the new basal body than the old, but as the cell progresses through the cell cycle (i.e. 2K2N) the *TbBB590::YFP* signal appears to be equal on both basal bodies. Following induction of *TbBB590* RNAi, *TbBB590::YFP* is rapidly lost from the cell but no obvious phenotype can be observed even after 96 hours (Fig 6.13-B-(2-5)). Loss of *TbBB590* does not have any effect on *TbRP2* localisation or expression.

**(A)**







**Figure 6.13** The effect of RNA interference-mediated ablation of *TbBBP590* on procyclic form cells. **(A)** Graph showing growth of *TbBBP590* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. **(B)** Non-induced and induced cells at 24h, 48 h, 72h and 96h post induction were harvested for immunofluorescence analysis. Cells were stained with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m.

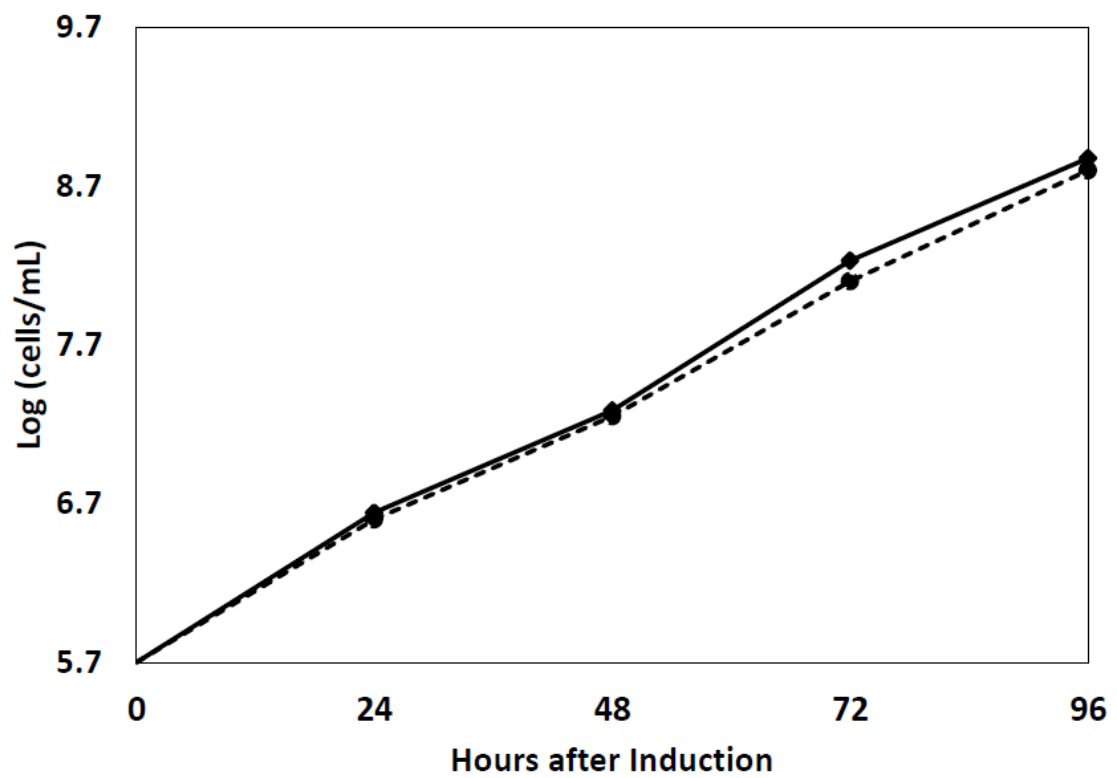
## 6.12 Tubulin polymerization promoting protein (TPPP/P25- $\alpha$ )

The identification of the tubulin polymerization promoting protein (TPPP/P25- $\alpha$ ) as a putative *TbRP2*-proximal protein was intriguing, given (i) the proposed role of TPPP in the polymerisation of tubulin into microtubules, and (ii) its restriction to eukaryotes that form cilia/flagella. To investigate the putative function(s) of the *T. brucei* *TbP25- $\alpha$*  orthologue, its expression was knocked by gene specific RNAi and the growth of RNAi induced and non-induced cells monitored every 24 hours (Fig 6.14-A).

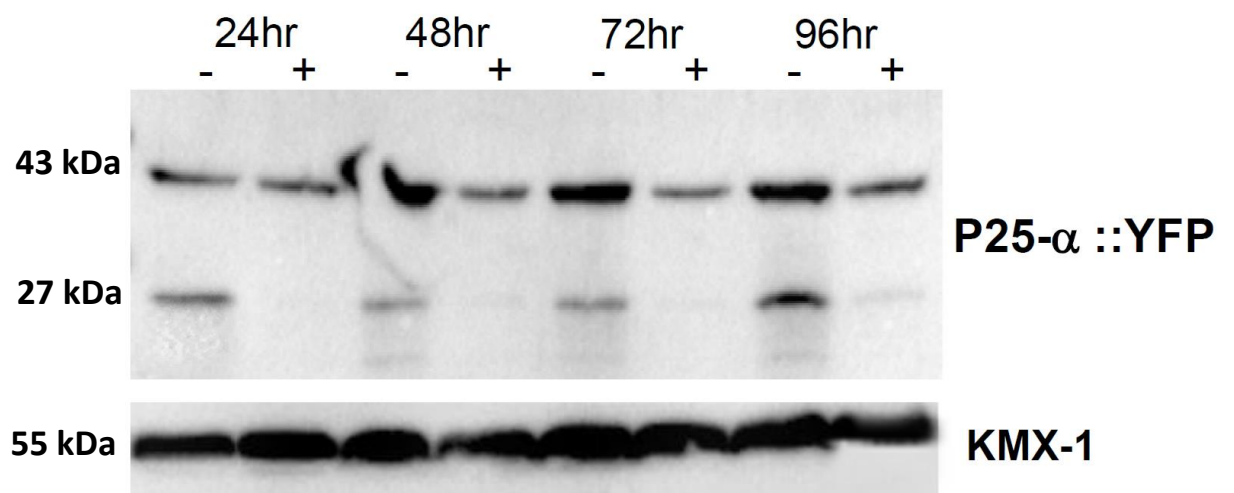
Based on this analysis, RNAi knockdown of *TbP25- $\alpha$*  has no effect on the growth and proliferation of procyclic form cells. RNAi knockdown of *TbP25- $\alpha$*  expression was analysed by immunoblotting using the anti-GFP antibody to detect *TbP25- $\alpha$ ::YFP* fusion protein; with KMX-1 detection of  $\beta$ -tubulin used as a loading control (Fig 6.14-B). In Fig 6.14-B, a dominant band of ~43 kDa (expected size of the *TbP25- $\alpha$ ::YFP* fusion protein) and a faint band of ~27 kDa (predicted size of YFP) was detected in the non-induced cells. The lower MW band is interpreted as being a cleavage product of the *TbP25- $\alpha$ ::YFP* fusion protein. After induction of *TbP25- $\alpha$*  RNAi, the expression of *TbP25- $\alpha$ ::YFP* was moderately reduced, however, *TbP25- $\alpha$ ::YFP* was still detectable after 96 hours of induction. Detection of  $\beta$ -tubulin by KMX-1 indicated that equal loading of non-induced and RNAi induced samples.

Fluorescence microscopy analysis reveals that in the absence of *TbP25- $\alpha$*  RNAi induction, *TbP25- $\alpha$ ::YFP* localises at cell body and the flagellum (Fig 6.14-C-1). However, after induction of *TbP25- $\alpha$*  RNAi, the expression of *TbP25- $\alpha$ ::YFP* was successfully depleted in some cells, while in others *TbP25- $\alpha$ ::YFP* expression appeared to be unaltered 96 hours after induction. No obvious phenotype was induced following loss of *TbP25- $\alpha$*  even after 96 hours (Fig 6.14-C), nor did loss of *TbP25- $\alpha$*  have any effect on *TbRP2* expression or localisation.

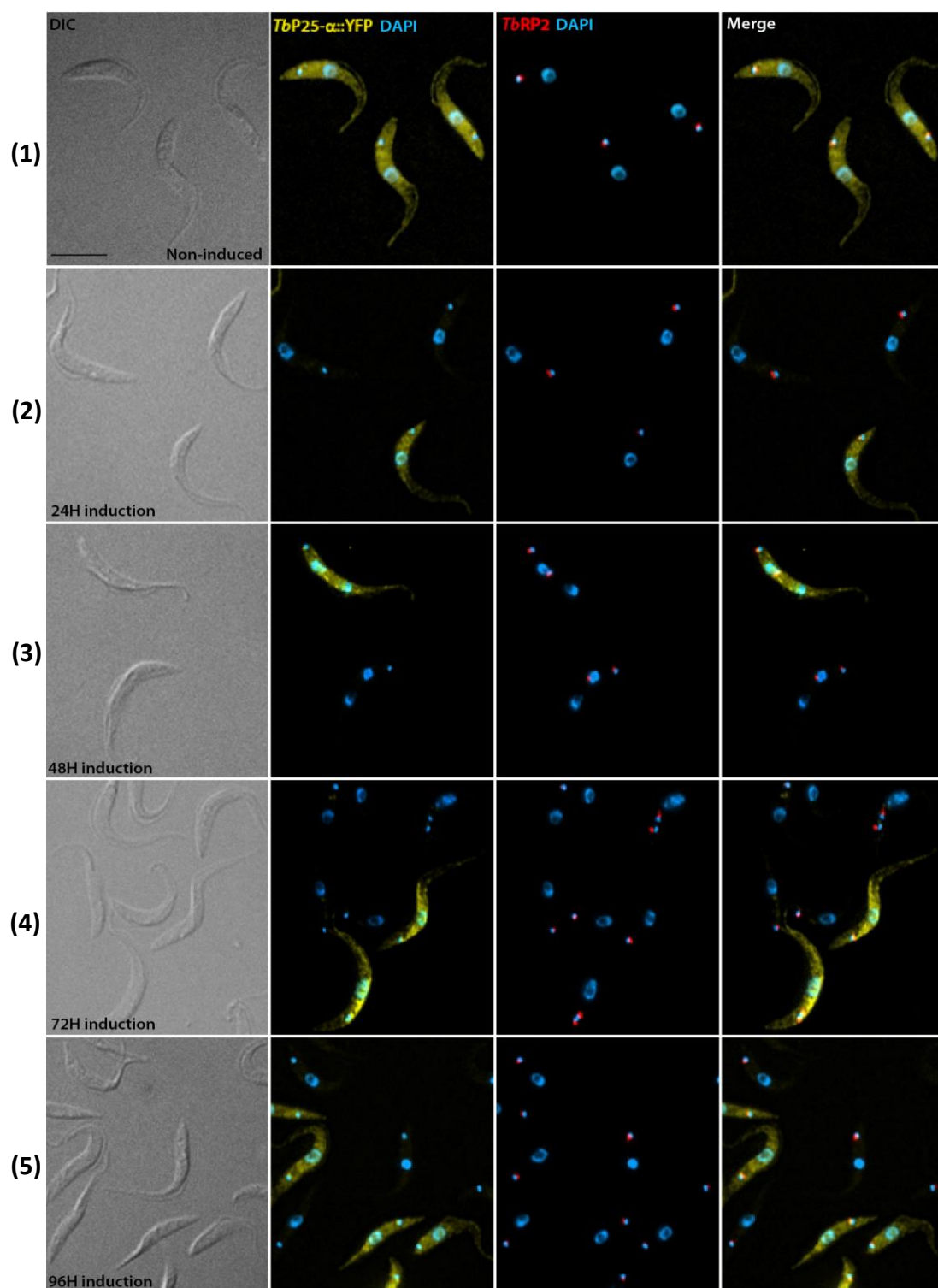
**(A)**



**(B)**



**(C)**



**Figure 6.14 The effect of RNA interference-mediated ablation of *TbP25-α* on procyclic form cells.** **(A)** Graph showing growth of *TbP25-α* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. **(B)** The *TbP25-α* expression level was monitored by western blotting. Both non-induced and induced cells were collected at 24h, 48 h, 72h and 96h post induction and probed with anti-GFP antibody. KMX1 detection of  $\beta$ -tubulin was included as a loading control. **(C)** Non-induced and induced cells at 24h, 48 h, 72h and 96h post induction were harvested for immunofluorescence analysis. Cells were stained with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m.

## 6.13 Effect of *TbRP2* RNAi on the expression and localisation of selected putative *TbRP2*-interacting/near neighbour proteins

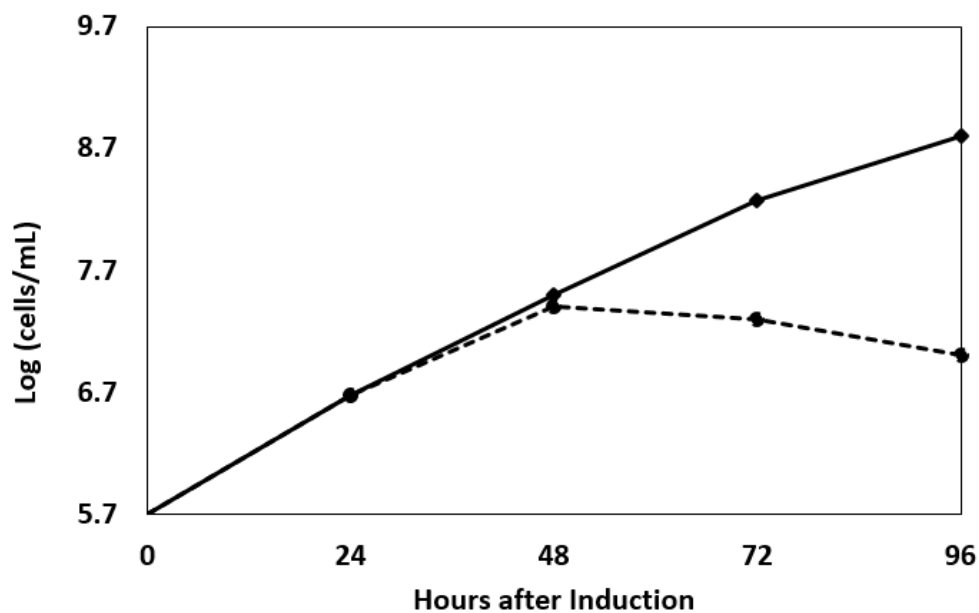
As RNAi mediated knockdown of all putative *TbRP2*-proximal proteins failed to perturb the basal body localisation of *TbRP2*, the reciprocal experiment was undertaken for a selected cohort i.e. the effect of *TbRP2* RNAi on the expression and localisation of KMP-11, HSP70, *TbBBP590*, ARL3 and BARTL1 was investigated. Cell lines capable of expressing YFP-tagged versions of KMP-11, HSP70, *TbBBP590*, ARL3 and BARTL1 were generated in a *TbRP2* RNAi (smOX P9) cell line. As described previously, following RNAi induction (in this instance *TbRP2* RNAi) cell growth was monitored and the expression and intracellular localisation of *TbRP2* and YFP-tagged proteins of interest investigated.

### 6.13.1 KMP-11

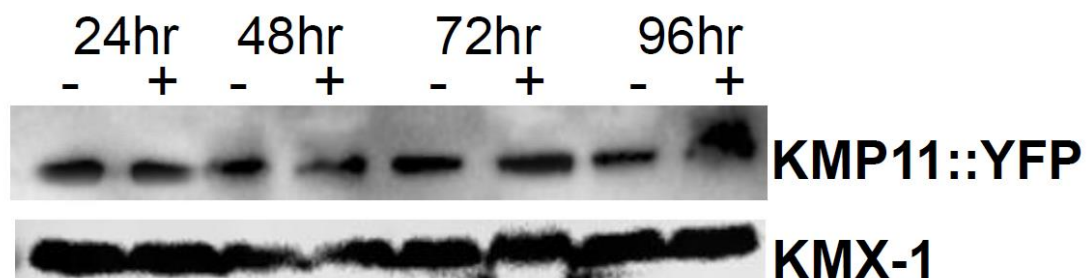
The induction of *TbRP2* RNAi results in a profound growth defect 48 hour after induction of gene-specific RNAi (Fig 6.15-A). Although *TbRP2* expression was not directly monitored in immunoblotting experiments, the flagellum assembly defect that arises following *TbRP2* depletion was apparent in RNAi induced populations (Fig 6.15-C-(1-4)). However, expression of *TbKMP-11::YFP* is clearly unaffected by induction of *TbRP2* RNAi-mediated ablation (Fig 6.15-B). Immunofluorescence analysis was also carried out; non-induced cells expressing *TbKMP-11::YFP* and co-labelled with the anti-*TbRP2* antibody are shown on the top panel (Fig 6.15-C-(1)). At 24 hours after induction of *TbRP2* RNAi the majority of cells look normal, *TbRP2* can still be detected by immunofluorescence analysis and there is no effect on *TbKMP-11::YFP* expression (Fig 6.15-C-(2)). At 48 hours post induction, cells in the culture flask were observed to be moving abnormally. When RNAi induction proceeds for 72 hours, more severe phenotypic abnormalities were observed in the population, the short flagellum formed in *TbRP2* RNAi induced cells demonstrated in cytokinesis defect resulting in an

accumulation of multi-nucleate cells (Fig 6.15-C-(4)). As described earlier, that *TbKMP-11* has multiple localisations within the PCF cells. When *TbRP2* is depleted PCF trypanosomes cells form an abnormally short new flagellum. The *T. brucei* flagellum is an essential and multifunctional organelle, which plays important roles in cell motility, but also in cellular morphogenesis and cell division. The consequence of forming a short new flagellum is the generation of multi-nucleated cells that have failed to divide correctly. In these cells, which have failed to assemble flagella, *TbKMP-11::YFP* accumulates within the cell body in discrete foci, potentially in the region of the kinetoplast.

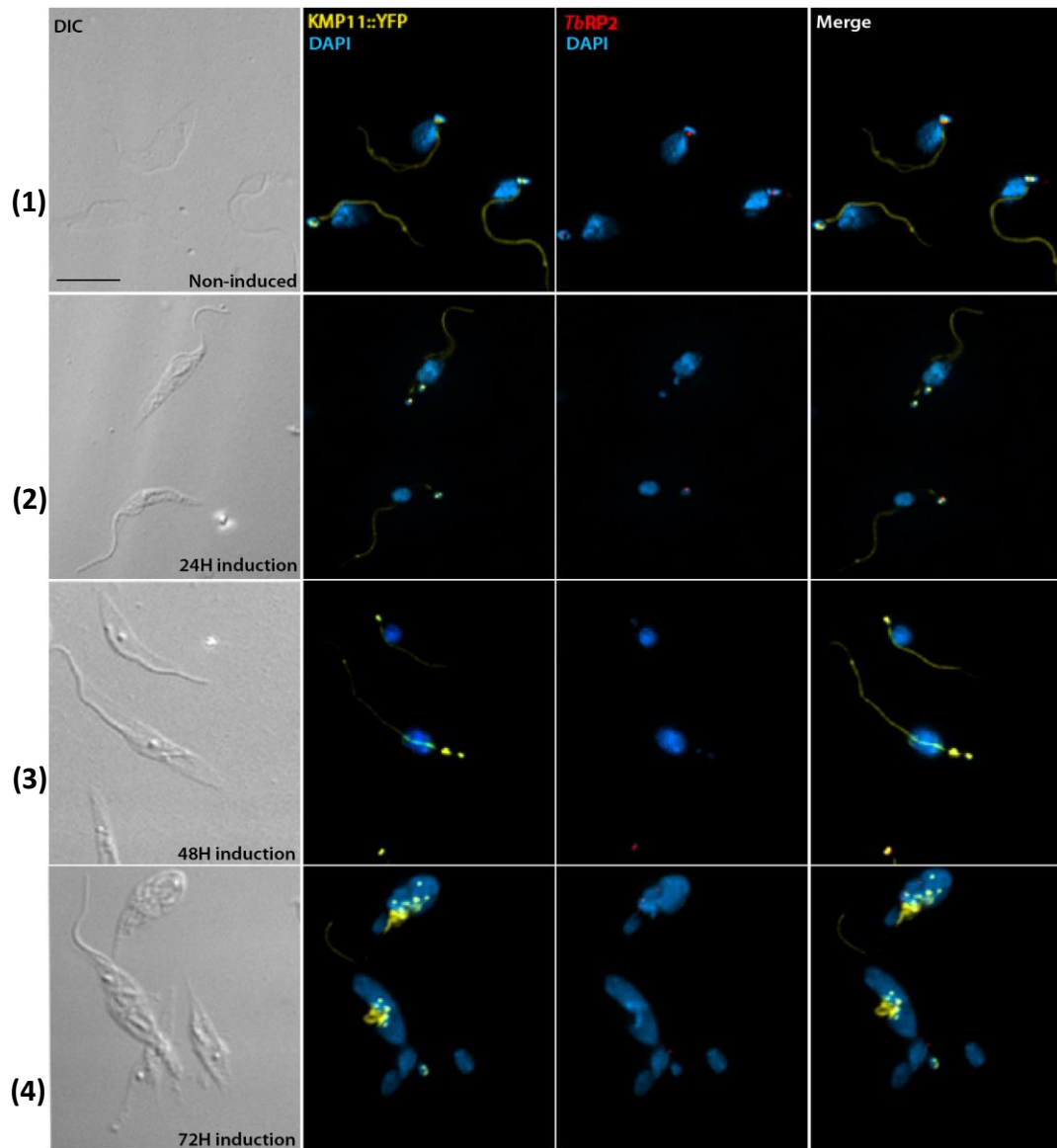
### (A)



### (B)



**(C)**



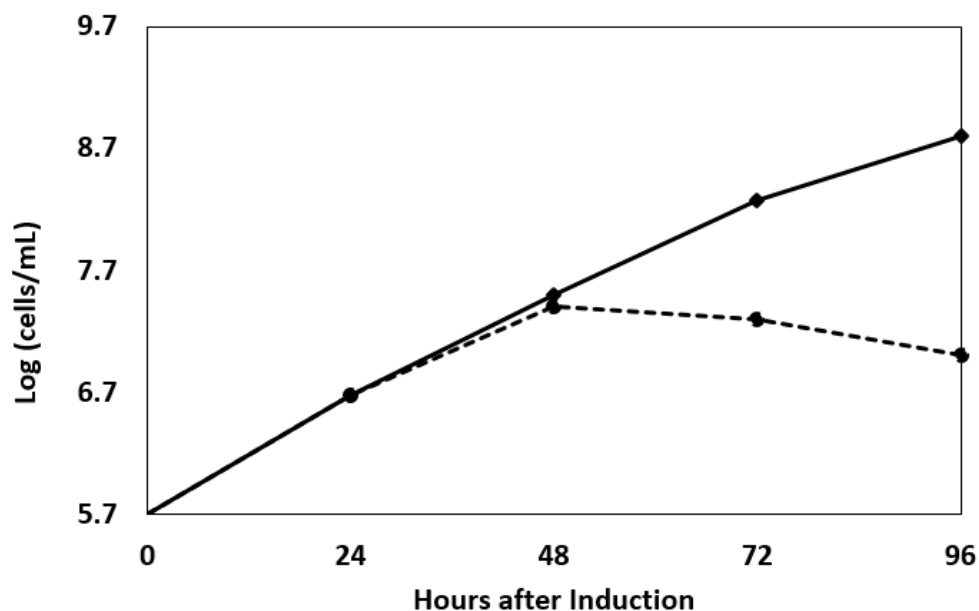
**Figure 6.15 The effect of *TbRP2* RNAi mediated ablation on *TbKMP-11* localization and expression. (A)** Graph showing growth of *TbRP2* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. **(B)** The expression of *TbKMP-11::YFP* on *TbRP2* RNAi background was monitored by western blotting. Both non-induced and induced cells were collected at 24h, 48 h, 72h and 96h post induction and probed with anti-GFP antibody. KMX1 detection of  $\beta$ -tubulin was included as a loading control. **(C)** Immunofluorescence images of *TbKMP-11::YFP*, *TbRP2* RNAi cell line. Cells at 24h, 48 h, 72h and 96h post induction were harvested, and co-labelled with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m.



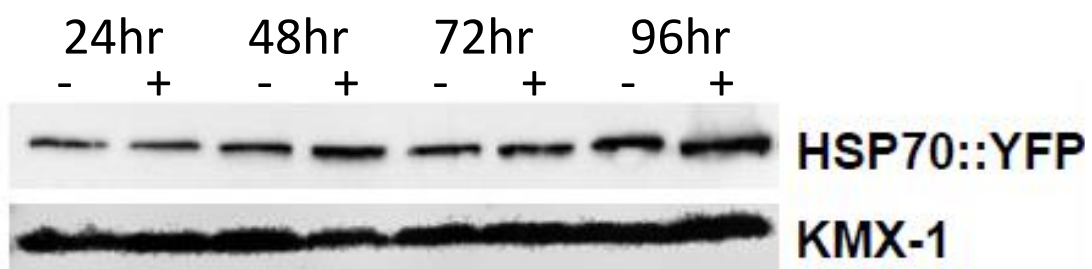
### 6.13.2 HSP70

The cell growth curve (Fig 6.16-A) shows that the *TbRP2* RNAi defect again became apparent 48 hours post RNAi induction, but immunoblotting experiments reveal *TbHSP70::YFP* expression is not affected by *TbRP2* depletion (Fig 6.16-B). Immunofluorescence analysis confirmed the loss of *TbRP2* expression at the mature basal body and the development of morphologically abnormal cells following *TbRP2* RNAi induction (Fig 6.16-C-(1-4)). Due to the loss of *TbRP2*, only accumulated *TbHSP70::YFP* signal can be seen within the shortened flagella (Fig 6.16-C white square), at later time point of RNAi induction (72H) (Fig 6.16-C white arrows).

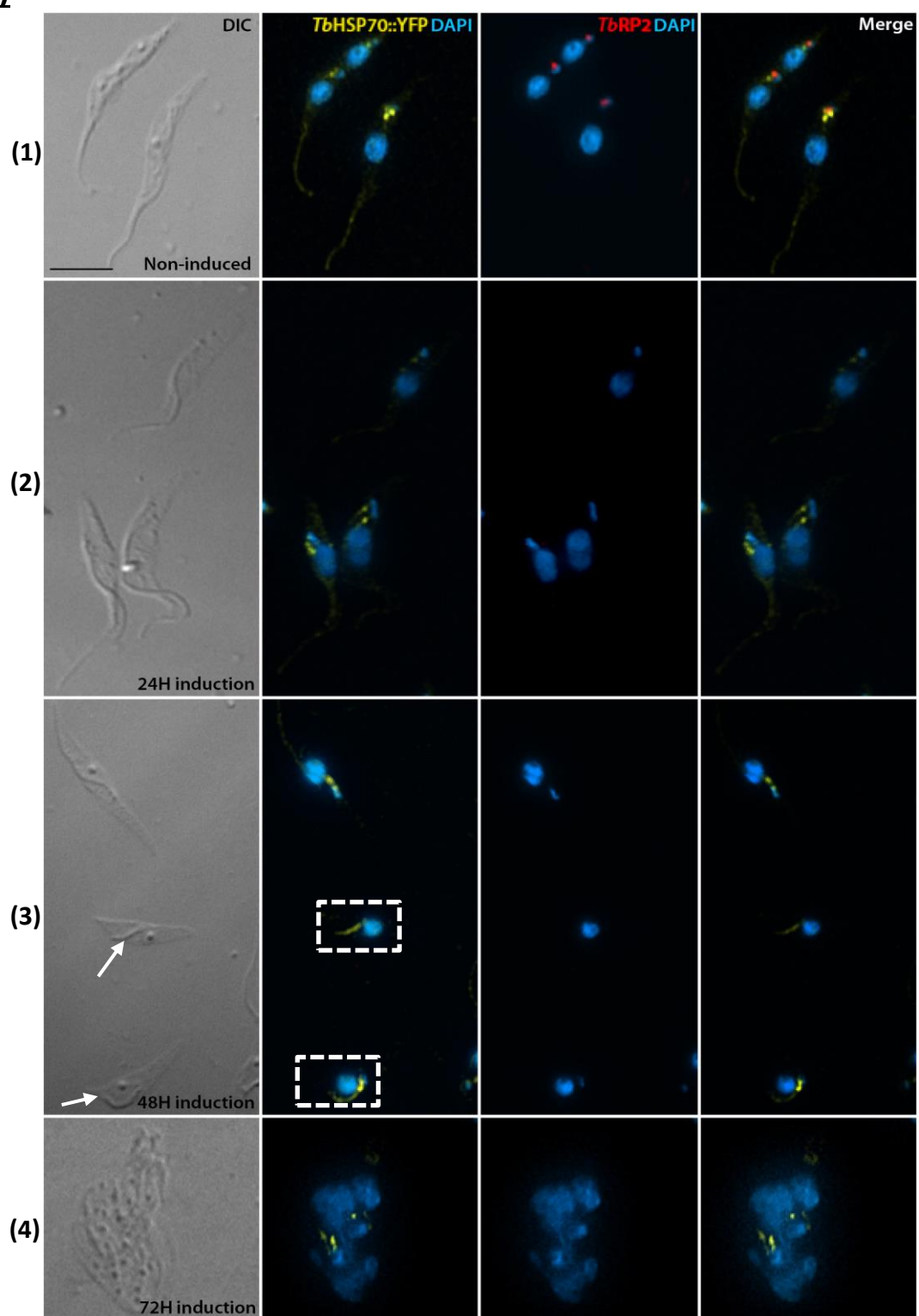
**(A)**



**(B)**



**(C)**



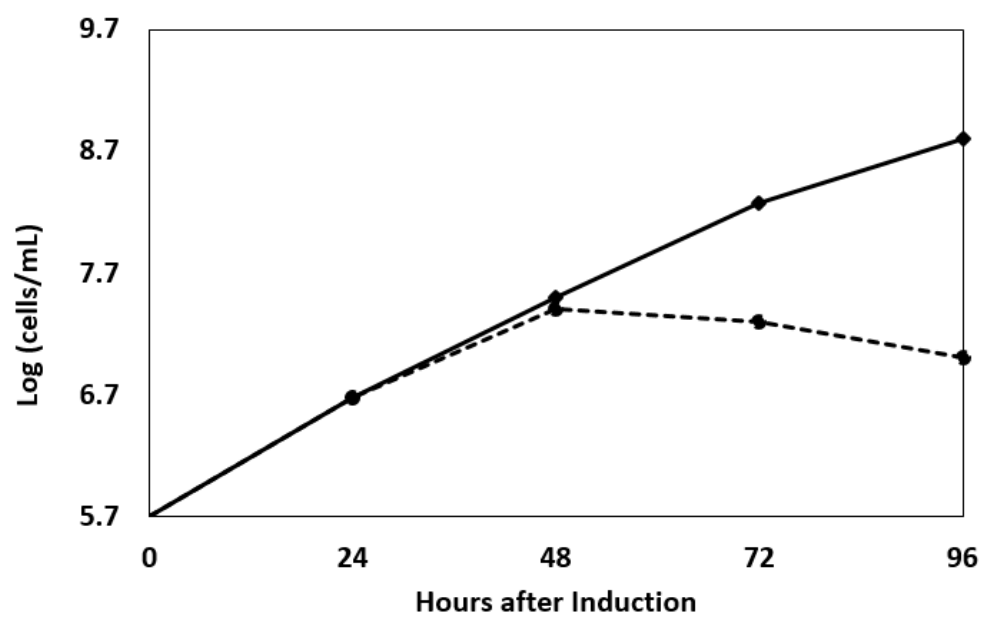
**Figure 6.16 The effect of *TbRP2* RNAi mediated ablation on *TbHSP70* localisation and expression.** **(A)** Graph showing growth of *TbRP2* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. **(B)** The expression of *TbHSP70::YFP* on *TbRP2* RNAi background was monitored by western blotting. Both non-induced and induced cells were collected at 24h, 48 h, 72h and 96h post induction and probed with anti-GFP antibody. KMX1 detection of  $\beta$ -tubulin was included as a loading control. **(C)** Immunofluorescence images of *TbHSP70::YFP*, *TbRP2* RNAi cell line. Cells at 24h, 48 h, 72h and 96h post induction were harvested, and co-labelled with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m. White arrow indicates short flagellum, and White square indicates accumulation of *TbHSP70::YFP* signal in shortened flagella

### 6.13.3 *TbBBP590*

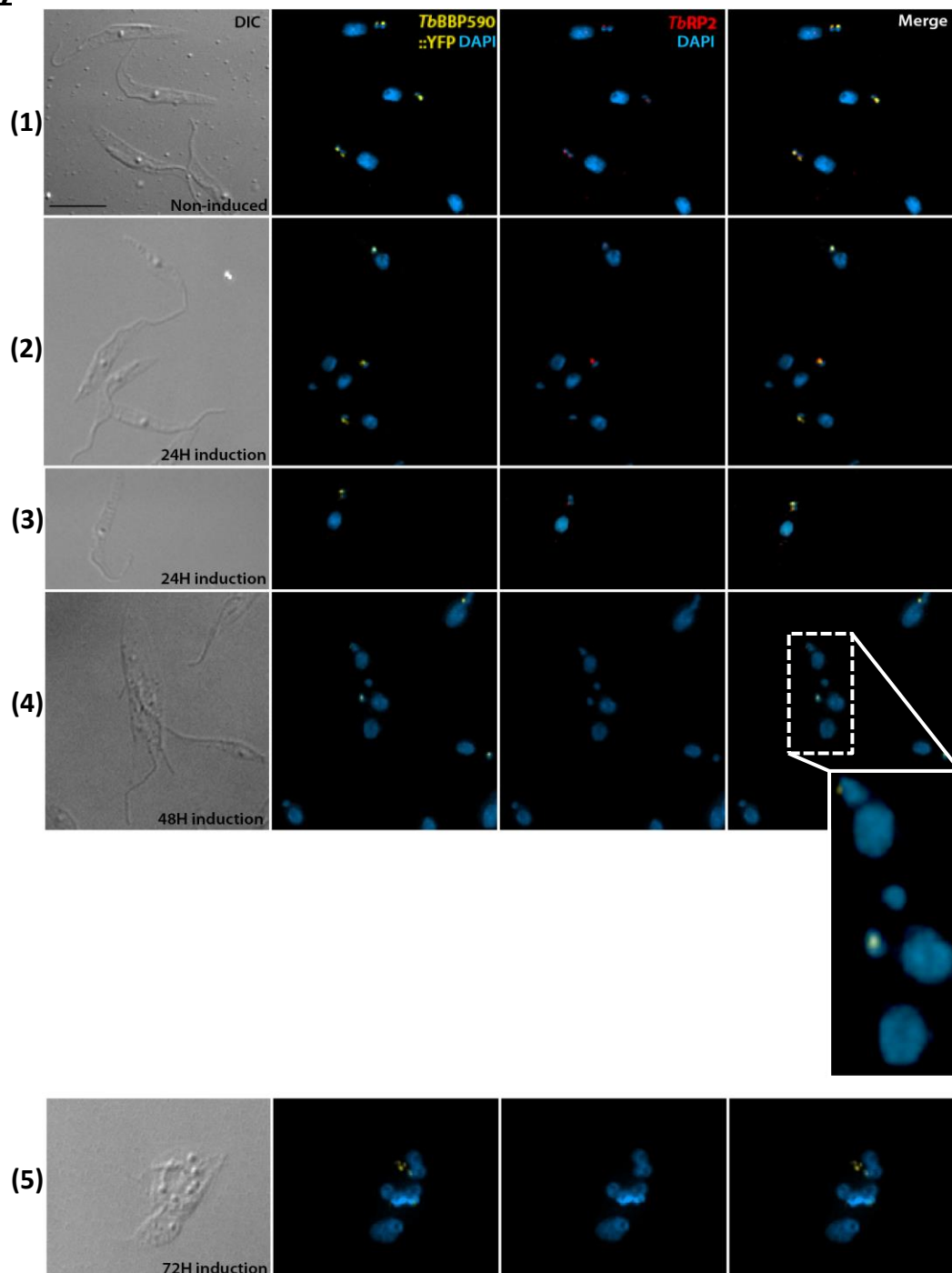
The cell growth curve again shows the *TbRP2* RNAi defect became apparent 48 hours post RNAi induction (Fig 6.17-A) and subsequent immunofluorescence analysis confirmed loss of *TbRP2* expression resulted in the development of morphologically abnormal cells (Fig 6.17-B). Due to the large MW of the *TbBBP590* protein it was not possible to investigate the expression of this protein by immunoblotting however, the expression of *TbBBP590::YFP* was monitored by fluorescence microscopy (Fig 6.17-B-(1-5)). The expected basal body localisation pattern of *TbBBP590::YFP* in non-induced cells is shown in the top panel of Fig 6.17-B-1.

In the presence of doxycycline, at 24 hours post induction, no obvious defect on cells can be observed (Fig 6.17-B-(2-3)). The expression of *TbRP2* was depleted in the majority of cells but there was no effect on the expression or localisation of the *TbBBP590::YFP* at that time. By 48 hours post-induction, visualisation of cells in culture revealed that cells had started to exhibit motility problems, and immunofluorescence analysis showed cells also displayed cell division defects (Fig 6.17-C-4; inset) leading to the formation of multi-nucleated cells. In this cell, which has three kinetoplasts, *TbBBP590::YFP* is associated with only one kinetoplast, but it is impossible to distinguish whether this is an 'old' or 'new' kinetoplast. When the RNAi induction time point reaches 72 hours, cells exhibiting more severe phenotypes were accumulating within the population; these were the consequence of cytokinesis defects arising from the expected short flagellum phenotype and the generation of multi-nucleated cells (Fig 6.17-C-5). Although *TbRP2* RNAi ablation might cause *TbBBP590* segregation problems in some cells, the *TbBBP590::YFP* signal can still be observed accumulating around the kinetoplast area in cells.

**(A)**



**(B)**

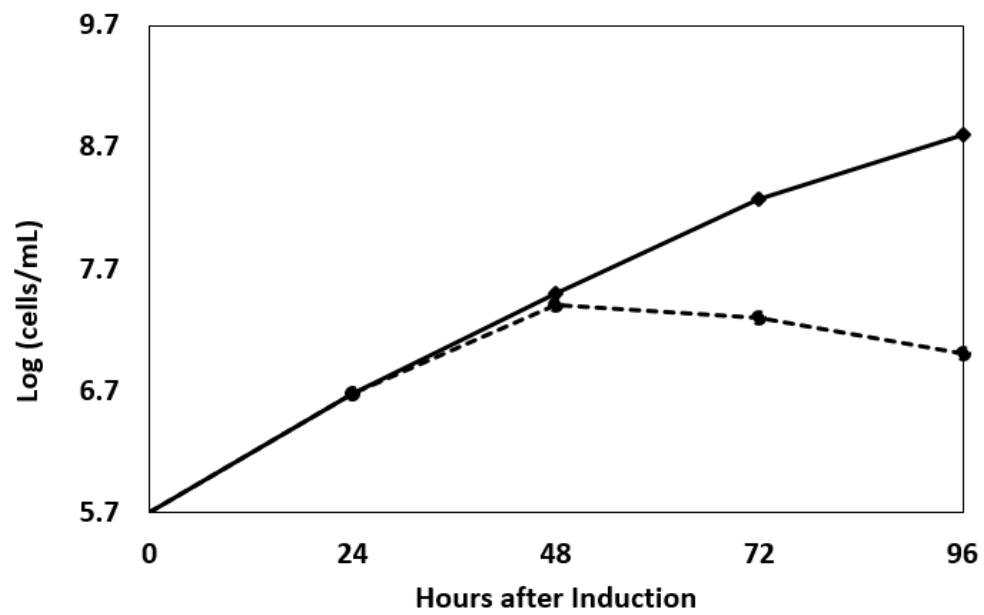


**Figure 6.17 The effect of *TbRP2* RNAi mediated ablation on *TbBBP590* localisation and expression. (A)** Graph showing growth of *TbRP2* RNAi induced (black dotted line) and non-induced (black solid line) procyclic trypanosome cells. **(B)** Immunofluorescence images of *TbBBP590::YFP* on *TbRP2* RNAi background. Cells at 24h, 48 h, 72h and 96h post induction were harvested, and co-labelled with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m. Inset indicates that the segregation of *TbBBP590::YFP* at the basal body might be affected by the depletion of *TbRP2* only in some cells.

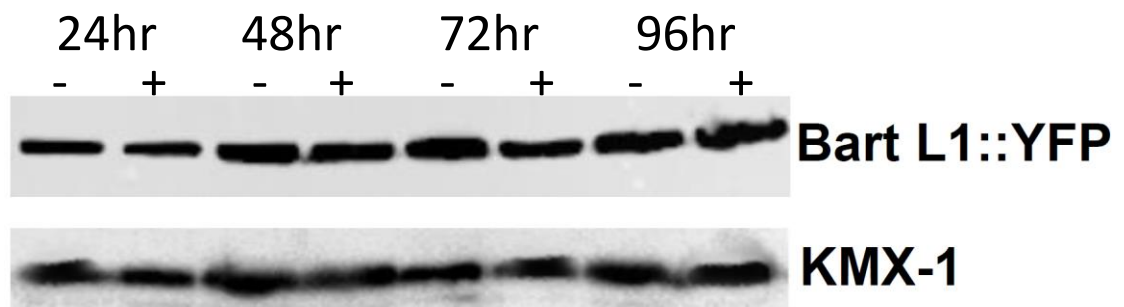
#### 6.13.4 BART LIKE 1

The induction of *TbRP2* RNAi results in a profound growth defect 48 hours after induction of gene-specific RNAi (Fig 6.18-A) but immunoblotting reveals that the expression level of *TbBARTL1::YFP* remains unchanged 96 hours after *TbRP2* RNAi induction (Fig 6.18-B). Thus *TbBARTL1::YFP* expression is not affected by the loss of *TbRP2* from cells. Subsequent immunofluorescence microscopy confirmed loss of *TbRP2* expression and the development of morphologically abnormal cells expected following *TbRP2* depletion (Fig 6.18-C-(1-5)). The localisation of *TbBARTL1::YFP* to the cell body and flagellum in non-induced cells is shown in the top panel (Fig 6.18-C-1); the cells were co-labelled with anti-*TbRP2* antibody to show the position of the mature basal body. After 24 hours of *TbRP2* RNAi induction, most cells have a normal morphology but *TbRP2* expression at the basal body is showing evidence of depletion. However, there is no effect on the expression of the *TbBARTL1::YFP* (Fig 6.18-C-2). At 48 hours post induction, cells in culture were moving abnormally slowly and when these cells were examined by immunofluorescence microscopy, *TbRP2* expression had been lost and cells presented with short flagella (Fig 6.18-C-3). After 72 hours of *TbRP2* RNAi induction, cells with anticipated cytokinesis defects and multi-nucleated cells accumulated within the population (Fig 6.18-C-4). However, there was no evidence that loss of *TbRP2* (and the generation of severe morphological abnormalities) affected *TbBARTL1* expression/localisation until 96 hours after induction (Fig 6.18-C-5).

**(A)**

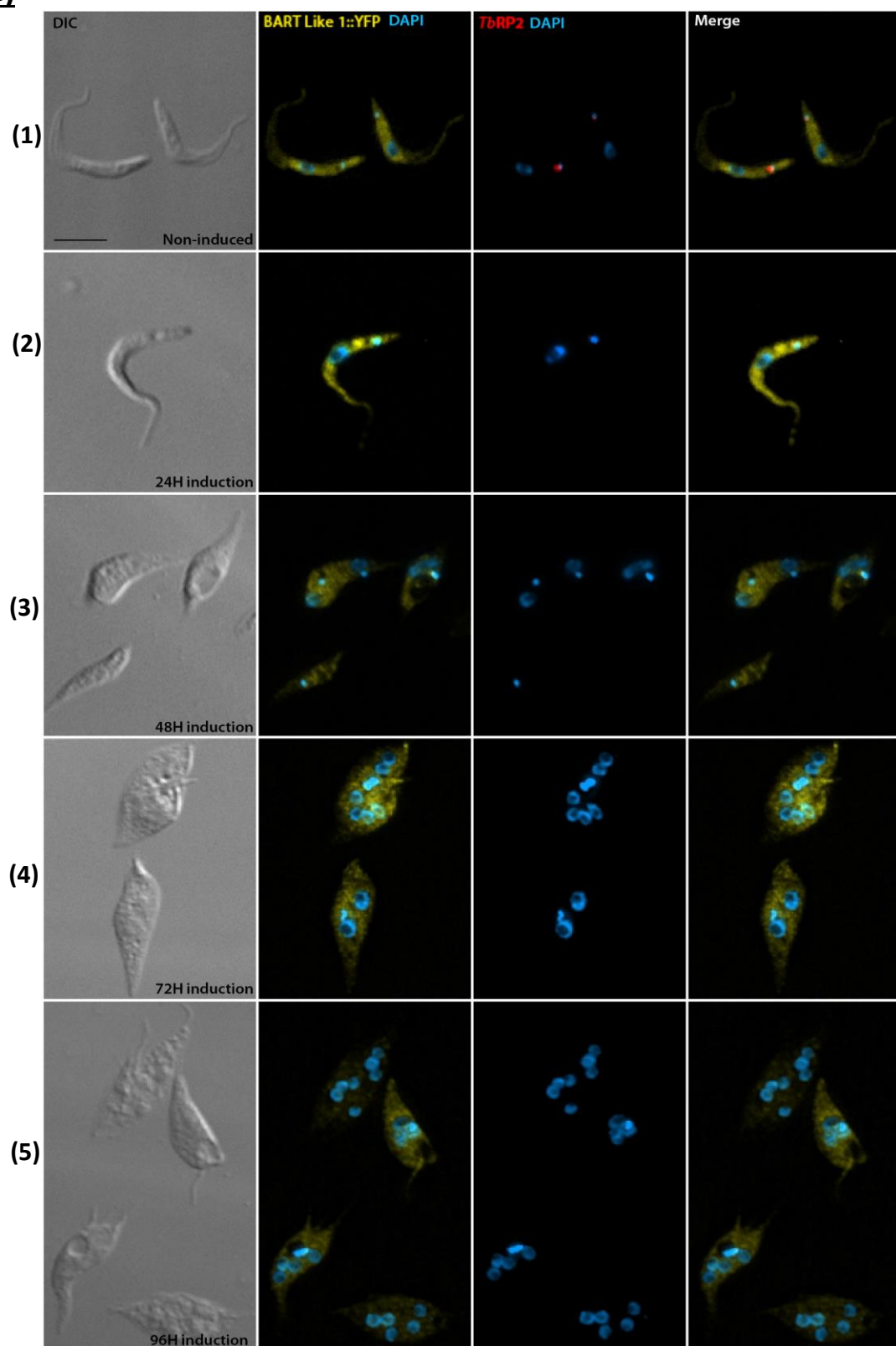


**(B)**





**(C)**

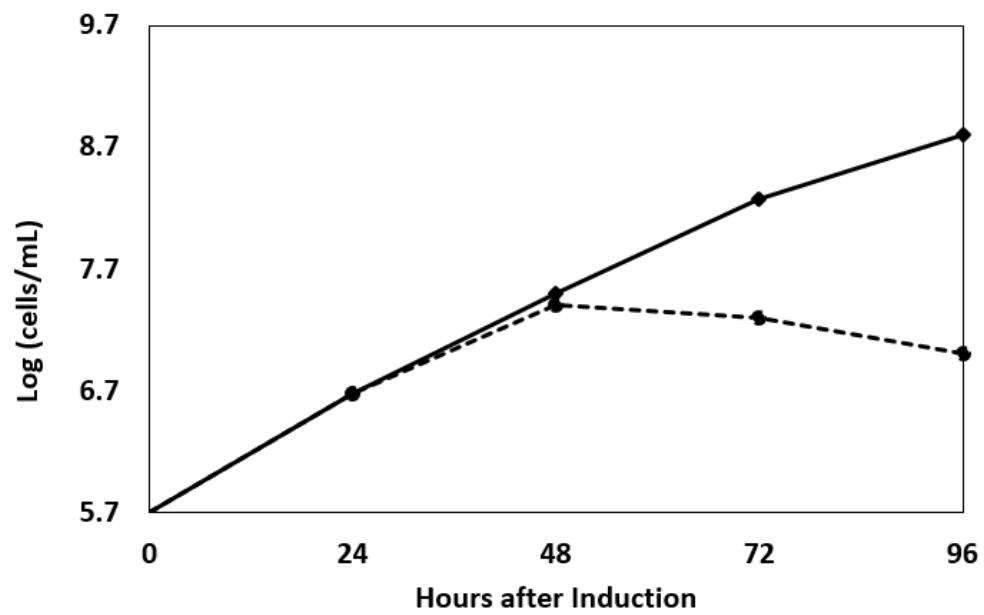


**Figure 6.18 The effect of *TbRP2* RNAi mediated ablation on *TbBARTL1* localisation and expression.** **(A)** Graph showing growth of *TbRP2* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. **(B)** The expression of *TbBARTL1::YFP* on *TbRP2* RNAi background was monitored by western blotting. Both non-induced and induced cells were collected at 24h, 48 h, 72h and 96h post induction and probed with anti-GFP antibody. KMX1 detection of  $\beta$ -tubulin was included as a loading control. **(C)** Immunofluorescence images of *TbBARTL1::YFP*, *TbRP2* RNAi cell line. Cells at 24h, 48 h, 72h and 96h post induction were harvested, and co-labelled with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m.

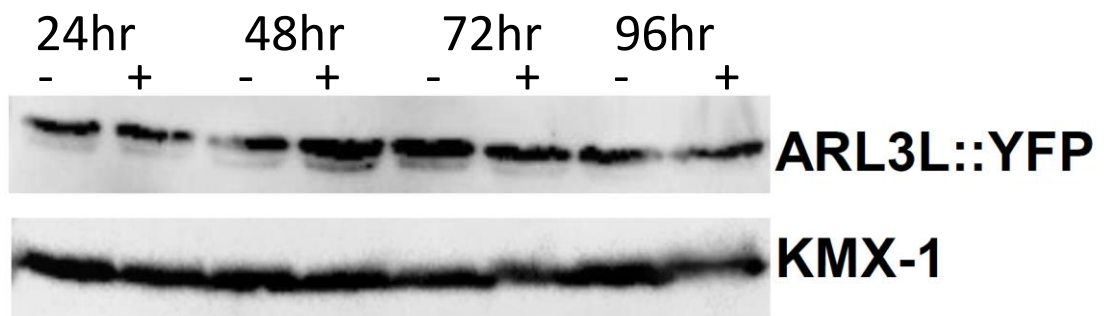
### 6.13.5 ARL3 Like

The induction of *TbRP2* RNAi results in a profound growth defect 48 hours after induction of gene-specific RNAi (Fig 6.19-A) but immunoblotting experiments revealed that the expression of *TbARL3L::YFP* remained unaltered 96 hours after induction *TbRP2* RNAi (Fig 6.19-B). Subsequent immunofluorescence microscopy confirmed loss of *TbRP2* expression and the development of morphologically abnormal cells expected following *TbRP2* depletion (Fig 6.19-C-(1-5)). The localisation of *TbARL3L::YFP* in non-induced cells are shown on the top panel of Fig 6.19-C-1; cells were co-labelled with the anti-*TbRP2* antibody, confirming the co-localisation of both proteins. Following induction of *TbRP2* RNAi, not only was the *TbRP2* protein lost from the basal body but intriguingly so was *TbARL3L::YFP*. This is most clearly observed 24 hours after *TbRP2* RNAi induction, where in one of the cells the old basal body is still *TbRP2* and *TbARL3L::YFP* positive, yet the basal body nucleating the new flagellum is negative for both proteins (Fig 6.19-C-2; see inset). At 48 hours post induction, the basal body localisation of both *TbRP2* and *TbARL3L::YFP* is abolished (Fig 6.19-C-3). However, *TbARL3L::YFP* expression is not lost entirely (confirming the immunoblotting result) as *TbARL3L::YFP* accumulated within cells with a punctate, yet dispersed, pattern. This result suggests that the basal body recruitment and/or retention of *TbARL3L::YFP* is dependent upon *TbRP2* expression. After 72/96 hours of RNAi induction, more severe phenotypes were evident within the population i.e. cells with cytokinesis defects and multi-nucleate cells with a short flagellum. In these cell types, *TbARL3L::YFP* is clearly evident within the cytoplasm (Fig 6.19-C-(4-5)).

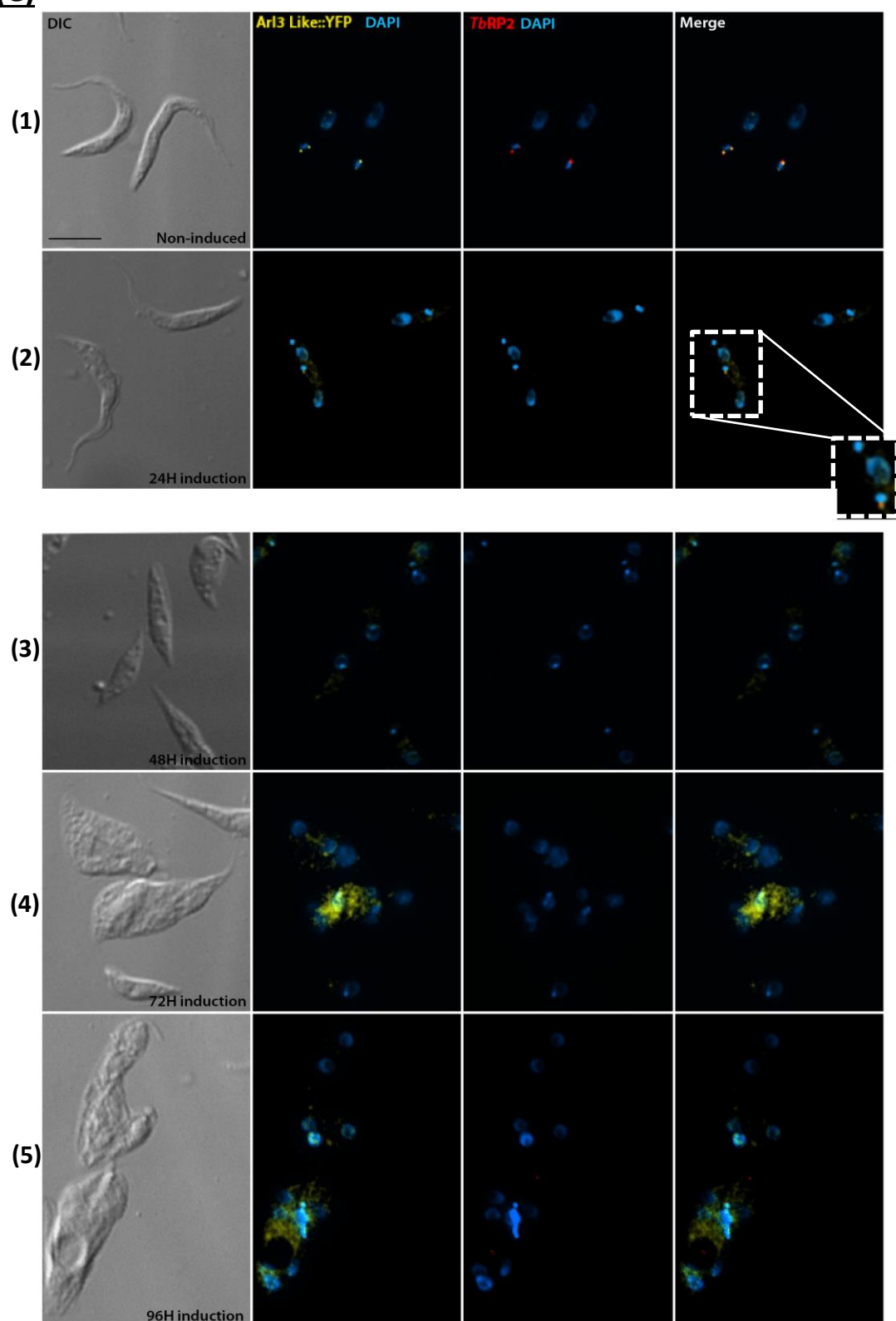
**(A)**



**(B)**



**(C)**



**Figure 6.19 The effect of *TbRP2* RNAi mediated ablation on *TbARL3L* localisation and expression.** **(A)** Graph showing growth of *TbRP2* RNAi induced (black broken line) and non-induced (black solid line) procyclic trypanosome cells. **(B)** The expression of *TbARL3L::YFP* on *TbRP2* RNAi background was monitored by western blotting. Both non-induced and induced cells were collected at 24h, 48 h, 72h and 96h post induction and probed with anti-GFP antibody. KMX1 detection of  $\beta$ -tubulin was included as a loading control. **(C)** Immunofluorescence images of *TbARL3L::YFP*, *TbRP2* RNAi cell line. Cells at 24h, 48 h, 72h and 96h post induction were harvested, and co-labelled with DAPI to label nuclei (N) and kinetoplasts (K) and with the anti-*TbRP2* antibody, which labels the mature basal body. Scale bar = 10  $\mu$ m.

## 6.14 Summary

This chapter describes the functional investigation of selected putative *TbRP2* interacting/nearby proteins by RNAi in PCF trypanosomes. Based on the bioinformatics analysis in Chapter 4, a total of 11 proteins were selected for analysis; 4 proteins are kinetoplastid-specific (*TbBBP590*, *TbCAP51*, *TbMARP* and *TbKMP-11*) and the other 7 have conserved sequences in other organisms (Summary Table 6.1). The localisation(s) of each protein are listed in the table below, as are the phenotypic consequences of RNAi ablation. Among these proteins, *TbKMP-11* (Li and Wang, 2008; Stebeck et al., 1995) and *TbCAP51* (Portman and Gull, 2014) have been studied previously and shown to be essential for cytokinesis in PCF trypanosomes. A sub-set of proteins were selected to investigate the phenotypic consequences of *TbRP2* RNAi ablation is summarised in Table 6.2.

Protein ID	Localisation(s)	RNAi (expression level knockdown?)	Phenotype(s)?	Effect on <i>TbRP2</i> ? (localisation /expression)
<b>ARL3 Like</b>	Basal body / nucleus	Yes	Growth defect after 72h (some abnormal cells)	No
<b>BARTL1</b>	Cell body / flagellum	Yes	No gross defects (a few cells show some abnormalities in cells)	No
<b>HSP70</b>	Cell body / flagellum	Expression came back after 96h induction	Cell growth defect between 24h – 48h but then recovers	No

<b>TCP-1-ε</b>	Cell body	Expression came back after 72h induction	Flagellum detachment and aberrant cell division	No
<b>VHS-domain Containing protein</b>	basal body, distal to BB	Yes	No apparent defect	No
<b>Protein Kinase</b>	basal body, distal to BB	Yes	No apparent defect	No
<b>CAP51</b>	Sub-pellicular microtubules	Expression came back after 72h induction	Aberrant cell division	No
<b>TbBBP590</b>	Basal body (odd pattern)	Yes	No apparent defect	No
<b>MARP</b>	Basal body	Yes	No apparent defect	No
<b>KMP-11</b>	Basal body, flagellum, flagellar connector, end of cell body	Yes	Flagellum detachment, multinucleate cells	No
<b>P25-α</b>	Cell body and flagellum	Only moderate Reduction	No apparent defect	No

**Table 6.1 Localisation and phenotypic consequences of RNAi ablation for proteins under investigation.**



Protein ID	<i>TbRP2</i> RNAi effect on YFP expression	Phenotype(s) <i>TbRP2</i> RNAi	Effect on	
			Localisation ?	Expression ?
<b>ARL3 Like</b>	No	Short and structurally aberrant flagellum	Yes (Cell body)	No
<b>BARTL1</b>	No	Short and structurally aberrant flagellum	No	No
<b>HSP70</b>	No	Short and structurally aberrant flagellum	No	No
<b><i>TbBBP590</i></b>	No	Short and structurally aberrant flagellum	No	No
<b>KMP-11</b>	No	Short and structurally aberrant flagellum	No	No

**Table 6.2 Summary of experiments investigating if loss of *TbRP2* affects the localisation and expression of a selected subset of the proteins identified by the *TbRP2* BioID experiments.**

## Chapter 7 Discussion

In my PhD study, I used BioID to identify candidate near neighbouring and/or interacting proteins of *TbRP2*. To give greater confidence to candidate proteins identified by Mass Spectrometry (MS), I also combined the BioID approach with SILAC methodology. In that regard, I note that another group recently combined SILAC with more conventional immunoprecipitation to identify with confidence novel components of the mitochondrial kinetoplast tripartite attachment complex (or TAC) (see section 1.3.2) (Käser et al., 2016). Bioinformatics analysis was subsequently carried out to prioritise which candidate *TbRP2* near-neighbour/interacting proteins would be subject to functional analysis on a basis of their evolutionarily conservation with trypanosomatids and other organisms. In Table 7.1, a summary of the work undertaken in this thesis in relation to the putative *TbRP2*-interacting proteins is presented. In addition, a discussion of what emerged from my BioID-based analysis of *TbRP2* function is summarised in this final chapter.

Protein ID	Localisation	RNAi phenotype(s)	RNAi knockdown efficiency	SILAC experiment (Cyto, WC or both)	Kinetoplastid Specific?	Evidence for functional interaction
<b>ARL3 Like</b>	Basal body /nucleus	Growth defect after 72h; enlarged flagellar pocket in small cell populations	Yes	WC	No	Yes
<b>BARTL1</b>	Cell body /flagellum	No gross defects (abnormalities in small number of cells)	Yes	WC	No	No
<b>HSP70</b>	basal body, distal to BB/flagellum	Cell growth defect between 24h – 48h but then recovers	Expression came back after 96h induction	Both	No	No
<b>TCP-1-ε</b>	Cell body	Flagellum detachment and aberrant cell division	Expression came back after 72h induction	WC	No	No
<b>VHS-domain Containing protein</b>	basal body, distal to BB	No apparent defect	Yes	WC	No	No
<b>Protein Kinase</b>	basal body, distal to BB	No apparent defect	Yes	Both	No	No

<b>CAP51</b>	Sub-pellicular microtubules	Aberrant cell division	Expression came back after 72h induction	Both	Yes	No
<b>TbBBP590</b>	Basal body (odd pattern)	No apparent defect	Yes	Cyto	Yes	No
<b>MARP</b>	Basal body	No apparent defect	Yes	Both	Yes	No
<b>KMP-11</b>	Basal body, flagellum, flagellar connector, end of cell body	Flagellum detachment, multinucleate cells	Yes	Both	Yes	No
<b>P25-<math>\alpha</math></b>	Cell body and flagellum	No apparent defect	Only moderate Reduction	WC	No	No

**Table 7.1 Summary table of all proteins selected for further studies in *T. brucei* following BioID analysis of *TbRP2::BirA\**. ‘Evidence for functional interaction’ refers to whether RNAi ablation experiments have provided evidence of functional interdependency between *TbRP2* and the stated protein.**

There was no evidence from either the SILAC ratio derived from MS experiments or by immunofluorescence analysis that axonemal tubulin is processed by *TbRP2*. Thus it is unlikely that *TbRP2* functions in quality control processing for tubulin as originally hypothesised (Stephan et al., 2007). It is more likely that *TbRP2* functions in more generalised protein trafficking capacity required for flagellum assembly. For example, I identified the likely ADP-ribosylation factor-like 3 (Arl3) related protein (Arl3L) with confidence as the interacting partner for RP2 GAP activity from MS and also from subsequent co-localisation and RNAi analyses. Following MS identification of *TbARL3L* as a candidate partner for *TbRP2*, immunofluorescence microscopy showed that *TbARL3L* was localised to the basal body, where it co-localised with *TbRP2*. *TbARL3L* also appeared to be present in the nucleus throughout the cell cycle although both signals were only seen on intact fixed cells (Fig 5.3). However, RNAi knockdown of *TbArl3L* did not lead to a shortening in flagellum length, as seen in *TbRP2* RNAi experiments, but in some cells caused flagellum detachment, cytokinesis defect and also overly enlarged flagellar pockets. The recruitment of *TbRP2* to the basal body was not dependent on *TbArl3L* expression, despite that they co-localised at the basal body (Fig 6.7). However, when *TbRP2* was knocked down the basal body localisation of *TbArl3L* was abolished and it relocated to the cytosol, providing the strongest evidence that *TbARL3L* is the partner protein for *TbRP2* GAP activity (Fig 6.19). In mammals there is plenty of evidence to indicate a variety of Arl-dependent complexes are associated with IFT and vesicular traffic necessary for cilium assembly and even in trypanosomatids additional Arl3-like proteins are involved in flagellum assembly and length control (Sahin et al., 2004). No additional Arl proteins, however, were identified from my BioID analysis. Fig 7.1 and Table 7.2 summarise current understanding of some of the Arl (Arf-like) and Rab proteins that have been shown participating in cilia formation and/or function in mammals (Li and Hu, 2011; Li et al., 2012).

Of all the small GTPase proteins, Arl6 is one of only two proteins that have been associated with ciliopathies (Fan et al., 2004). Genetic mapping and comparative genomic analysis identified that *ARL6* was the causal locus for *BBS3* (Bardet-Biedl syndrome (BBS)) (Chiang et al., 2004). Later crystal structure analysis confirmed

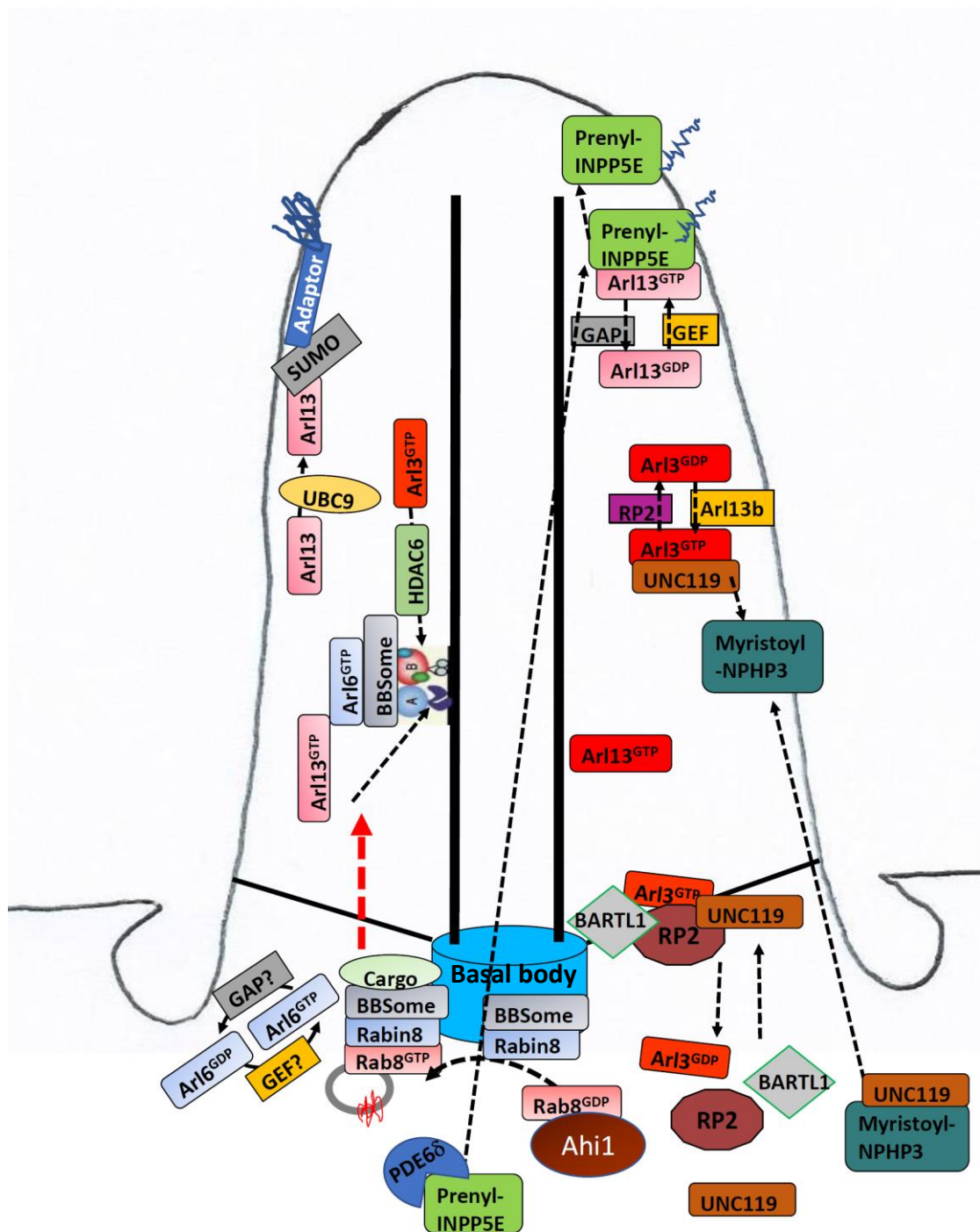
mutations around the ARL6 GTP binding motif were present in patients with BBS; suggesting the importance role of small GTPase activity in ciliopathies (Wiens et al., 2010). In ciliated IMCD3 cells, Arl6 localised at the distal end of the basal body. Mutations in Arl6 reduce the length and number of primary cilia, and abolish its function in the Wnt signalling pathway (Wiens et al., 2010). In the mouse model, Arl6 is not essential for BBSome formation, but is required for localising melanin concentrating hormone receptor 1 (MCHR1) to the ciliary membrane and somatostatin receptor 3 to the entire cilium (Jin et al., 2010; Zhang et al., 2011). Additionally, Arl6 may also be involved in the SHH signalling pathway, as loss of Arl6 can disturb the retrograde transport of the sensory receptor of Smoothened (Zhang et al., 2011).

Arl13b was the second protein identified in patients with the classical form of Joubert Syndrome (JBTS). JBTS is characterised by polydactyly, cystic kidney, blindness, and abnormalities of the CNS (Cantagrel et al., 2008). *Arl13b* null mouse embryos also exhibit typical ciliopathic phenotypes, such as defects in neural tube patterning, limbs and eyes that mimic the JBTS phenotype (Caspary et al., 2007). Originally, Arl13b was identified in zebrafish as a highly-enriched cilium protein and critical for cilia formation in multiple organs (Duldulao et al., 2009; Sun et al., 2004). Ultra-structurally, cilia in both *arl-13*<sup>-/-</sup> worms and *Arl13b*<sup>hnn</sup> mice, are highly similar (i.e. truncated axoneme plus defective B-tubule closing) indicating an evolutionarily conserved function ARL-13/Arl13b in cilia (Caspary et al., 2007). In 2012, Humbert and colleagues identified a protein-protein interaction network that consists of ARL13B, phosphodiesterase 6δ (PDE6δ), centrosomal protein 164 (CEP164) and inositol polyphosphate-5-phosphatase E (INPP5E) in mammalian cells (Humbert et al., 2012). INPP5E is localised to primary cilia through a motif near the C terminus and prenyl-binding protein phosphodiesterase 6D (PDE6D)-dependent mechanisms and mutations in this gene can be found in JBTS patients (Bielas et al., 2009; Jacoby et al., 2009). Loss of Inpp5e activity leads to organ developmental defects such as cerebral developmental, skeletal defects, cystic kidney etc. (Bielas et al., 2009; Jacoby et al., 2009). In humans, Arl13b (not ARL2 or ARL3) is responsible for INPP5E targeting to the cilia, and mutations in Arl13b abolish the connection between ARL13B and INPP5E. This finding provides us

with addition information on proteins complexed at the base of the cilium that are linked to ciliopathies (Humbert et al., 2012).

Rab GTPases are the largest family of small GTPases, and their function and distinct localisation at intracellular membranes have been studied extensively. However, the association of several Rab proteins with ciliary trafficking and aberrant functions in ciliopathies has only recently been appreciated (Stenmark, 2009). So far Rab8, Rab10, Rab11, Rab17 and Rab23 have been shown to be involved in ciliary trafficking (Hsiao et al., 2012). One of the well-known functions of Rab8 is involvement in exocytic targeting in epithelial cells and neurons, as well as being a key player in rhodopsin transport and maintenance of the outer segment disk structure of rod cells (Deretic et al., 1995; Huber et al., 1993). Nachury and colleagues demonstrated that Rab8 GDP/GTP is functionally linked to the BBS complex, which localises to the basal body (Nachury et al., 2007). Later, other studies suggested that Rab8 may not function alone, but rather its activity is co-ordinated with other Rab proteins, such as Rab11 (Knödler et al., 2010). These workers demonstrated that Rab11-GTP-bound form interacts with Rabin8 directly, which stimulates the GEF activity of Rabin8 toward Rab8. Loss of Rab11, by RNAi or mutation, blocks primary ciliogenesis. This highlights the importance of Rab GTPases and their coordinate action with each other in the regulation of vesicular trafficking during primary ciliogenesis (Knödler et al., 2010).

Rab23 was first identified as the gene mutated in the open-brain mouse (Eggenchwiler et al., 2001), and acts as a negative regulator of Shh signalling. Rab23 localises to the plasma membrane, endosomes, as well as the cytosol. In humans, mutations in the *Rab23* gene cause Carpenter's syndrome, which is characteristic by obesity, craniosynostosis, polydactyly, and soft-tissue syndactyl (Jenkins et al., 2007). By using FRAP (fluorescence recovery after photobleaching), Boehlke and co-workers showed that when Rab23 is depleted it alters the steady state levels of Smo in the cilium, but have no effect on the microtubular tip protein EB1 or the receptor protein Kim1. Based on their proposed model, the function of Rab23 is to mediate the ciliary turnover of Smo, and may possibly function in a recycling pathway (Boehlke et al., 2010).



**Fig 7.1. Working model of the role of Arls and Rabs in the cilium.** Arl13-GTP enhance the interaction between IFT A and B complexes. The SUMOylation of Arl13 by UBC19 is required for ciliary targeting. The BBSome recruitment is needed by Arl6 and move into cilium cross the ciliary gate. RP2 acts as a GAP for Arl3, which release the NPHP3 into the cilium, and stabilise IFT-B and KIF17 motors associations and trafficking. BB-basal body. Adapted from (Zhang et al., 2013)



Name	GAP/GEF	Effector(s)	Functions
<b>ARL3</b>	XPR2/ARL13b	HDAC-6 ( <i>C. elegans</i> )  UNC119, CCDC104,  Golgin-245, BART, C20orf194 and PDE6δ	<ul style="list-style-type: none"> <li>• Negative regulator for Ciliogenesis</li> <li>• Signalling pathway</li> </ul>
<b>ARL6</b>	?	The BBSome (human)	<ul style="list-style-type: none"> <li>• Membrane trafficking</li> <li>• Regulating BBSome</li> </ul>
<b>ARL13b</b>	?	UBC-9 ( <i>C. elegans</i> )  INPP5E (human)	<ul style="list-style-type: none"> <li>• Ciliogenesis by regulating IFT complex</li> </ul>
<b>Rab8A</b>	?/RPGR, RABIN8	myosin Vb	<ul style="list-style-type: none"> <li>• Cilia formation</li> <li>• Regulating ciliary membrane sorting and trafficking</li> </ul>
<b>Rab10</b>	?	SEC16A;  GDI1-2	<ul style="list-style-type: none"> <li>• Membrane transport</li> </ul>
<b>Rab11</b>	?	myosin Vb	<ul style="list-style-type: none"> <li>• Upstream for Rab8</li> <li>• Ciliary membrane assembly</li> </ul>
<b>Rab23</b>	Evi5-L/?	patched 1	<ul style="list-style-type: none"> <li>• Regulating hedgehog pathway</li> </ul>
<b>Rab28</b>	?	?	<ul style="list-style-type: none"> <li>• Protein trafficking in photoreceptor cells</li> </ul>

**Table 7.2. Some of the ciliary Arls, Rabs and their possible interactors.**

I draw attention to the identification of BARTL1 in Figure 7.1; an orthologous *T. brucei* protein was identified in my BioID analysis albeit without the confident SILAC ratio seen for other candidate near neighbour/interacting proteins to *TbRP2* (Fig 7.1). Despite the unfavourable SILAC ratios I also subjected *TbBARTL1* to experimental analysis. During the course of my study, Lokaj and colleagues identified BARTL1

localised to the transition zone in mouse inner medullary collecting duct (IMCD3) cells, act as a binding partner for Arl3 and GTP-bound form of Arl2. The presence of the BARTL1 protein enhances the interaction between RP2 and Arl3-GTP, which promoted the release of carrier proteins such as PDE $\delta$  and Unc119a/b (Lokaj et al., 2015). Therefore, even though the SILAC ratio for BARTL1 was not above the threshold I had selected, further analysis was carried out on this protein. In mammalian cells, BARTL1 both biochemical and structural analysis shows that the BART-like domain can recognise and interact with a LLxLLxxL motif within an N-terminal amphipathic helix of Arl3, which might suggest BARTL1 is involved in the Arl3 transport network in cilia. However, RNAi ablation of either BARTL1 or Arl3 had no effect on the localisations of the reciprocal protein in mouse renal epithelial cells (IMCD3). The explanation for this was suggested as relating to the low efficacy of siRNA. In *T. brucei*, BARTL1 protein was linked to suramin efficacy and potential resistance (Alsford et al., 2012). In this study, immunofluorescence microscopy showed that *TbBARTL1* is localised to the cell body, as well as into flagellum throughout the cell cycle (Fig 5.7). Again, the majority of cells in the population did not appear to present with any overt morphological phenotype when *TbBARTL1* was subjected to RNAi knockdown. However, I did observe that a small number of cells appeared to have some problems; e.g. multinucleated cells with segregation and cell cytokinesis defects, and also detached flagella. Whether the phenotype I observed was caused by loss of *TbBARTL1* or for other reasons is not clear at present. However, *TbRP2* expression and localisation was not dependent on *TbBARTL1* (Fig 6.8).

Of the proteins identified as candidates near neighbour/interacting of *TbRP2*, none of the ciliary gate components that have previously been shown to be required for cilium/flagellum assembly (see section 1.2.2) were present in my data sets. Instead, what emerged is the possibility that *TbRP2* functions alongside a variety of chaperone and other accessory proteins. For instance, proteins I tagged in my study included HSP70, TCP-1- $\epsilon$ , VHS and p25- $\alpha$ . HSP70 is probably one of the most well studied and most conserved proteins in evolution (Daugaard et al., 2007), and like most chaperones plays an important role in catalysing folding (Mayer and Bukau, 2005).

Moreover, HSP70 proteins have also been associated in ciliary protein transport (Bhogaraju et al., 2013; Kampinga et al., 2009). As early as 1995, Bloch and Johnson identified a putative flagellar Hsp70 in the green algae *C. reinhardtii* (Bloch and Johnson, 1995). They showed the localisation of HSP70 in the flagellum is enriched at the distal end; i.e. the site of axonemal assembly *in vivo*. Axonemal-bound HSP70 was released when the cells were treated under ionic conditions. Together this may suggest HSP70 chaperone targets tubulin to the flagellar tip, where it participates in axonemal assembly (Bloch and Johnson, 1995). In the BHK21 (kidney) cell line, HSP70 was purified from membrane-bound organelles (MBOs), and can release cargo from kinesin-driven anterograde IFT in an Mg-ATP-dependent and *N*-ethylmaleimide-sensitive manner, which might indicate HSP70 functions by releasing kinesin from cargo in specific subcellular domains to effect flagellar protein transport (Tsai et al., 2000). Later studies carried out by Shapiro and colleagues confirmed the co-localisation of HSP70A and the FLA10 (IFT kinesin-II) by immunofluorescence analysis, and they additionally proposed that instead of acting as part of the IFT machinery HSP70 was cargo carried by IFT (Shapiro et al., 2005). More recently, a study suggested that the function of HSP70 in the cilium/flagellum might be to co-operate with other ciliary proteins. IFT88 belongs to IFT B complex and is essential of ciliogenesis. By using Yeast two-hybrid and pulldown assays *in vitro*, IFT88 was shown to directly interact with a mammalian relative of DNAJ (MRJ) protein. Evidence was also shown that MRJ is associated with specific membrane protein guanylyl cyclase 1 (GC1) in retinal photoreceptor cells and that RNAi knockdown caused a reduction of GC1::GFP in the cilium. Furthermore, MRJ/HSP70 chaperone pair are co-immunoprecipitated with KIF3A and IFT88. It is possible therefore that MRJ/HSP70 enhances the interaction between membrane proteins GC1 and IFT88 in ciliary transport (Bhowmick et al., 2009).

Based on genome annotation, there are 12 HSP70 proteins in *T. brucei* (Folgueira and Requena, 2007) (see section 4.1). In this study, the localisation of *Tb*HSP70 was in part at basal body, but also can be found in flagellum as well, and so the question is whether this *Tb*HSP70 is functioning as a chaperone alone, or whether it works

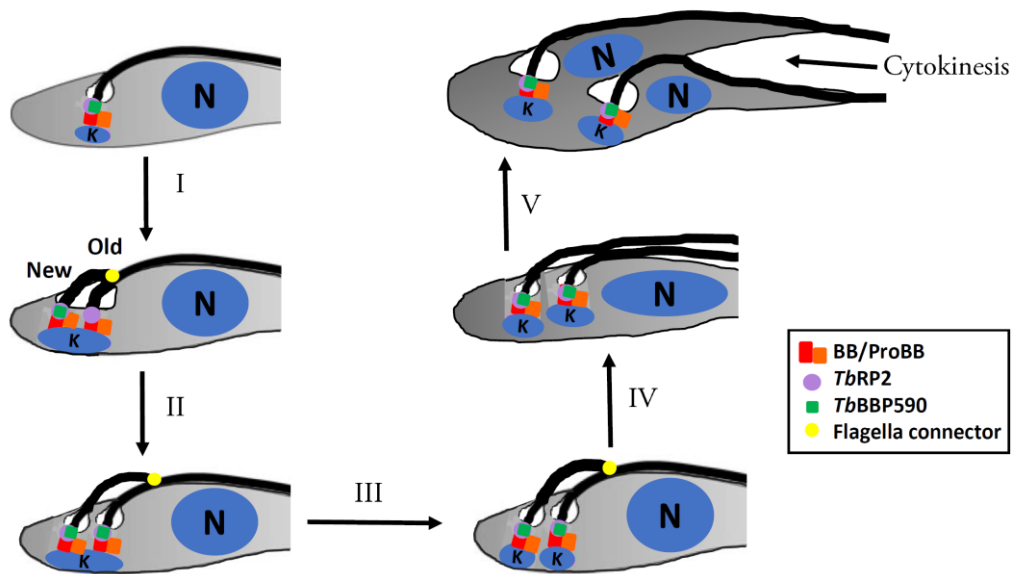
together with other heat shock proteins to function in the cilium/flagellum. It has been widely accepted that HSP70 interacts with HSP40 to stimulate HSP70 ATPase activity to bind nearby client proteins (Schlecht et al., 2011). However, I draw attention to the fact that in *Chlamydomonas* both HSP40 and HSP70 are found in the flagella (Bloch and Johnson, 1995; Yang et al., 2005). In fact, HSP40 is a component of the radial spoke complex in *Chlamydomonas*, and in sperm flagella of the ascidian *Ciona intestinalis* (Satouh et al., 2005). Depletion of RSP16 (a member of HSP40 family) results in normal length flagella, but flagella movement is affected, resulting in a twitching flagellum phenotype (Yang et al., 2008). In mammalian cells, heat shock factor-1 (HSF-1) plays a critical role in maintaining cilia beating and possibly by regulating the expression of HSP90 during tubulin polymerisation (Takaki et al., 2007). Inhibition of HSP90 leads to a reduction in the average length and beat frequency of primary cilia. In addition, when the cells were treated with elevated temperatures (from 37°C to 42°C) reveals that the frequency and length of cilia were shortened further, which might indicate a possible role for HSP90 in the maintenance of primary cilia (Prodromou et al., 2012).

Other chaperones which involved in cilia function are three chaperonin-like BBS proteins BBS6, BBS10, and BBS12 (not components of the BBSome) form a protein complex with other six CCT chaperonins (BBS/CCT complex) localised to the centrosomes and is required for BBSome assembly (Seo et al., 2010). BBS6, BBS10, and BBS12 are all share sequence homology to the CCT/TRiC family of group II chaperonins (Katsanis et al., 2000; Stoetzel et al., 2006; Stoetzel et al., 2007). Although those three proteins form a complex with CCT chaperonins, but their function do not involve in protein folding with ATPase activity, as they lack the ATP binding motif to server the purpose. The possible roles of those chaperones could be involved in preserving ciliary proteins' native functional conformation during and after ciliogenesis, i.e. participating in the assembly or maintenance of large ciliary protein complexes. Interestingly, neither BBSome nor IFT were identified as near-neighbours of *TbRP2* and given the volume of traffic through the TFs one might expect some degree of biotinylation. This might indicate the *TbRP2* and other identified proteins as a distinct protein complex at the base of the flagellum.

I also identified a variety of cytoskeletal proteins such as *TbCAP51*, *TbBBP590* and *TbMARP*. *TbMARP* was the first protein identified in my BioID-SILAC experiments. Although in later experiments the SILAC ratio for this protein fell into the 'background noise', as this protein was recognised with large molecular weight and highly competitive repeated (Schneider et al., 1988) it was of interest as it could function in supporting basal body architecture at the base of flagellum. The localisation of YFP::*TbMARP* on whole cells was revealed as the subpellicular MT corset, with an additional signal distal to the basal body possibly being associated with an endosomal compartment (Fig 5.15). However, a pool of *TbMARP* at the basal body became more obvious when I analysed detergent-extracted cells and isolated flagella (Fig 5.16-5.17). These observations speak to the growing evidence that many proteins are multifunctional in cells, and are often found in multiple locations (Ginger, 2014). The identification of *TbCAP51* from the BioID experiments was not expected, as this protein has subpellicular corset localisation and the SILAC ratio suggests it is a good candidate. During the time of my study, Portman and Gull also identified *TbCAP51*, along with another protein *TbCAP51V* (*Tb927.7.2650*), as being localised to the subpellicular microtubule cytoskeleton, but the two proteins are subject to stage-specific expression in PCF and BSF respectively. *TbCAP51* plays an essential role in PCFs, RNAi ablation leads to a cytokinesis defect and results in the accumulation of multi-nucleated cells and/or cells with abnormal numbers of kinetoplasts, and a large number of zoids (i.e. 1K0N cells) (Fig 6.5). Interestingly, in another study of basal body proteins - *TbSAS-4* (which also localises to the tip of the FAZ), *TbCAP51* was also identified by BioID, but no further comment was made on the reason why *TbCAP51* had been identified in this study (Hu et al., 2015).

From my BioID candidates, there was another high molecular weight protein (encoded by *Tb927.11.10660*) with a predicted highly repetitive secondary structure, which localised to the basal body. During the time of my PhD, Dang and co-workers also identified this protein and named it *TbBBP590*, from BioID studies carried out using *TbCEP57* as bait. *TbCEP57* has been shown to localise between the mature and probasal body, and is essential for basal body separation (Dang et al., 2017). In their

study, a single *TbBBP590* signal was found near the mature basal body (Dang et al., 2017). However, the distinct temporal pattern of *TbBBP590* observed in this study (Fig 6.18-6.19) was not mentioned in Dang et al., (2017). At the beginning of the cell division cycle, the trypanosome cell possesses a single basal body, from where the single flagellum nucleates, and an immature (pro-basal) body lying immediately alongside. This stage is normally called 1K1N, and these cells display a single focus of *TbBBP590* at the mature BB that co-localises with *TbRP2*. As the cell division cycle progresses, the kinetoplast starts to replicate and the pro-basal body elongates to mature into a basal body and nucleates a new flagellum; these cells have two foci of *TbRP2*. Curiously, I observed that the focus of *TbBBP590* on the old basal body is lost at this stage (or is below the level of detection by DeltaVision microscopy), but is retained on the new basal body. However, by the G2 phase of the cell division cycle the 'old' basal body reacquires *TbBBP590* expression and both 'old' and 'new' basal bodies then remain *TbBBP590* positive until the end of the cell cycle (see Fig 7.2). Based on my RNAi analysis, loss of *TbBB590* does not lead to any obvious phenotype in procyclic form cells after 96 hours of induction. Loss of *TbBB590* also has no effect on *TbRP2* localisation or expression at the mature basal body (Fig 6.13). The spatial and temporal pattern of expression shown here for *TbBB590* is, to my knowledge, unprecedented in *T. brucei*. However, this unusual inheritance pattern is unlikely to be due to selective proteolytic cleavage of the YFP tagged *TbBB590* protein, as both the N- and C-terminally tagged *TbBB590* proteins show the same inheritance pattern on both cytoskeletons and isolated flagella. But, at present the function and thus the reason for loss and reacquisition is unclear.



**Figure 7.2 Schematic diagram of the recruitment of *TbBBP590* in *T. brucei* procyclic (insect) form cells during cell division cycle. (I)** G1 cells possess a single kinetoplast and nucleus (termed 1K1N), as well as one mBB (where the flagellum nucleated) and a ProBB (1 focus *TbRP2* and 1 focus *TbBBP590* co-localising at the mBB). The maturation and elongation of the ProBB in to a basal body, and new flagellum emerging from the flagellum pocket (2 foci *TbRP2* on both new and old mBB, but the focus *TbBBP590* can only be seen on the new mBB, not on the old). **(II)-(IV)** as the cell cycle progresses, the accumulation of *TbBBP590*::YFP focus reappear on the old mBB again.

KMP-11 can be found in wide variety of kinetoplastids (Stebeck et al., 1995), however, the KMP-11 localisation of varies between species. Two independent studies have shown that KMP-11 is localised at the basal body and flagellum, and that it plays an essential role in cytokinesis and is critical for kinetoplast/basal body segregation in PCF and BSF cells (Li and Wang, 2008; Zhou et al., 2015). RNAi ablation of KMP-11 also disturbs the FAZ filament in PCF trypanosomes (Zhou et al., 2015). Zhou and colleagues also indicated that KMP-11 endogenously tagged with a triple HA tag also localised to the FAZ filament in detergent-extracted cytoskeletons. KMP-11::3HA protein signal overlapped with CC2D (Zhou et al., 2011) at the FAZ in immunofluorescence experiments (i.e. two distinct KMP-11 signals were evident on the flagellum and FAZ). Moreover, the study confirmed that depletion of KMP-11 disturbs the stability of the FAZ and causes flagellar detachment (Zhou et al., 2015).

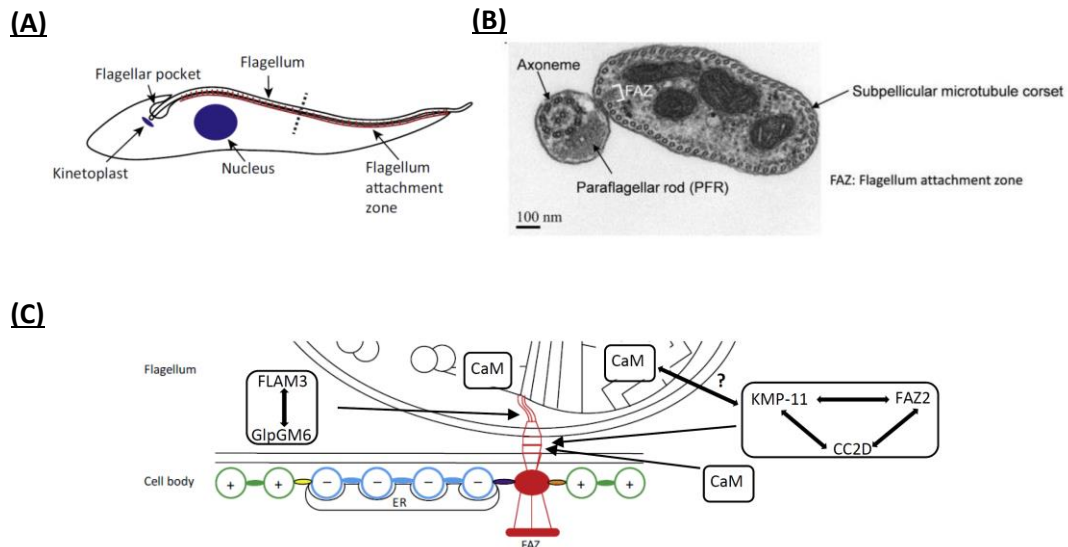
Based on my work, additional localisations of KMP-11 have been identified which were not mentioned in previous studies (Fig 5.4-5.6). Firstly, KMP-11::YFP does not exactly co-localise with *TbRP2* at the mature basal body, but is positioned more proximal, and there is always a “two-dot” pattern of KMP-11 coincident with *TbRP2*. Interestingly after 24 hours of KMP-11 RNAi induction, KMP-11 was barely detectable in the flagellum, but the KMP-11::YFP signal near the kinetoplast was still visible. Considering that the KMP-11 protein is positioned between the DAPI stained kinetoplast- and the basal body protein *TbRP2* (at least in this study), might suggested that KMP-11 plays a role in connecting the kinetoplast and the basal body (e.g. part of the tripartite attachment complex). As RNAi knockdown of KMP-11 resulted in defective kinetoplast segregation (Fig 6.1).

I also noticed that the localisation pattern of flagellum of KMP-11 and FLAM3 were similar (Sunter et al., 2015). FLAM3 was first identified as a new FAZ protein involved in flagellum adhesion in PCF (Rotureau et al., 2014). However, later studies showed that the flagellum detachment phenotype could be due to the mis-interpretation of FAZ reduction, and a longer unattached flagellum in PCF (Sunter et al., 2015). FLAM3-deletion leads to a reduction in FAZ length and causes morphological changes in cells, similar to the phenotype showed by the RNAi depletion of the calpain-like protein ClpGM6 (Hayes et al., 2014). Intriguing, KMP-11 RNAi ablation only inhibits new FAZ formation, which lead to detachment of new flagellum in the PCF, but not in BSF (Li and Wang, 2008). The FAZ has been proposed to control the position and polarity of the cleavage plane (Kohl et al., 2003), and is required for flagellum attachment and cytokinesis in PCF (Li and Wang, 2008; Vaughan et al., 2008). Together, the data may indicate that BSF and PCF trypanosomes may have different mechanisms of cytokinetic initiation. However, knockdown of FLAM3 does not affect KMP11 localisation or expression (Fig 6.2). This might indicate that KMP-11 (flagellum to cell body) is linked via FLAM3 (flagellum side) and so might explain why loss KMP-11 results in more severe aberrant phenotypes than FLAM3 RNAi.

In the KMP-11 localisation chapter, I mentioned that when cells were co-labelled with the anti-PFR2 antibody L8C4 (Fig 5.5), the KMP-11::YFP signal extended beyond the



L8C4 signal towards the posterior end of the cell. This indicated that KMP-11 is likely to be associated with the axoneme rather than the PFR because the PFR attaches to the axoneme only after it emerges from the flagellar pocket (Koyfman et al., 2011). Many PFR proteins are required for maintaining normal cell motility; for example PFR1 and PFR2 (the two main protein components of the PFR) are required for normal motility (Bastin and Gull, 1999; Bastin et al., 1998; Hughes et al., 2012; Koyfman et al., 2011). Calmodulin (CaM) RNAi also disturbs PFR assembly and leads to a reduction in the FAZ, but has no obvious effect on axonemal microtubules. Flagellum detachment is due to the fact that cells lose the FAZ connection to the cell body. KMP-11 localisation and expression remains unaffected around the kinetoplast/basal body area, however, there is a 'blob-like' accumulation of KMP-11 signal at the distal end of the flagellum and/or end of the cell body (Fig 5.3). This accumulation can also be seen in other PFR RNAi mutants (Bastin et al., 1999). One possible explanation could be that KMP-11 protein is transported by IFT to the end of cell body but due to a PFR assembly defect (e.g. as seen following CaM-depletion) KMP-11 cannot be delivered beyond. However, it remains unclear whether a KMP11-containing filament equivalent to the FAZ filament is formed in flagellum along the adhesion zone and whether FLAM3 physically interacts with CaM or any other protein(s) such as FLAM3. Nevertheless, it is very clear that the KMP-11 protein plays important roles in maintaining the flagellum to cell-body connection which might require protein-protein interactions from both the flagellum and cell body (see summary figure 7.3 below).



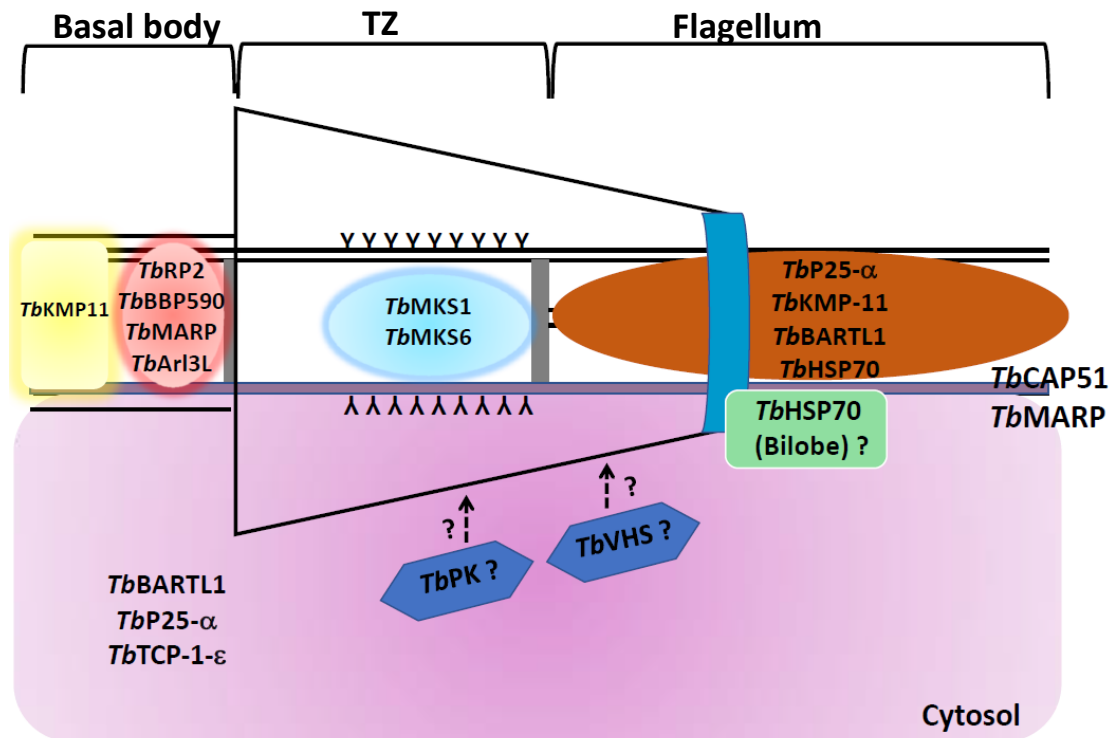
**Figure 7.3 Potential interactions and localisations of selected FAZ proteins.** (A) Schematic illustration of a trypanosome cell showing the relative locations of its single-copied organelles. (B) Cross-section of a trypanosome cell visualised via transmission EM. (C) Proteins within the black circle have been experimentally studied, and “?” indicates possible interaction. (A) and (C) adapted from (Sunter and Gull, 2016) and B taken from (Zhou et al., 2014)

A possible working model of how *TbRP2* and the putative *TbRP2*-interacting/proximal proteins identified by my BioID study may interact at the basal body is summarised in Fig 7.4. RNAi-mediated ablation of most of the putative *TbRP2* interacting proteins had no effect on *TbRP2* expression/localisation (with the possible exception of *TbCAP51* and *TbTCP-1-ε*). This indicates recruitment and retention *TbRP2* at the basal body does not depend on any of these proteins.

Based on my DeltaVision microscopy analysis, *TbPK* and *TbVHS* do not appear to co-localise with *TbRP2*, but both are positioned close to *TbRP2* at the basal body; which might indicate localisation to the endomembrane system. Bioinformatics analysis reveals that *Tb*protein kinase (Tb927.9.6560) share a 30% of amino acid identity to human Adaptor-Associated Kinase 1 (AAK1) protein (AAI43711.1). Like mammalian GAK family members, AAKs have been shown associated clathrin-mediated trafficking processes and essential for *T. brucei* endocytosis (Allen et al., 2003), and act as a positive regulator of Notch pathway (Gupta-Rossi et al., 2011). Orthologues of this protein kinase are well conserved across different organisms and in other closely related trypanosomatids, which might suggest this protein has a conserved role in

membrane-trafficking system. In a most recent study, this protein Kinase has been named as *TbAAKL*, and is a clathrin-associated protein, which is not essential for flagellum assembly and maintaining endocytic transport (Manna et al., 2017).

Immunofluorescence microscopy showed that *TbVHS::YFP* was localised to the basal body, and also distal to the basal body in a region that could be endosomal (Fig 5.10-5.11). In eukaryotic organisms, endosomes are positioned at such crucial area as numbers of trafficking pathways are passing from there, such as endocytic, phagocytic, Golgi to lysosome trafficking etc., which ensures that effectors and cargo proteins correctly reach their final destination (Jovic et al., 2010). Several experimental and structural studies have indicated that different trafficking pathways passing through the endosomes are coordinately regulated by proteins belongs to ENTH (Epsin NH<sub>2</sub>-terminal homology) / ANTH (AP180 N-terminal homology) / VHS-domain containing proteins (De Craene et al., 2012; Ford et al., 2001; Hyman et al., 2000; Itoh and De Camilli, 2006; Maldonado-Báez and Wendland, 2006; Mao et al., 2000). Interestingly, crystal structure analysis reveals that VHS and ENTH domain are very similar (Hyman et al., 2000; Mao et al., 2001). In *T. brucei*, the ENTH-domain protein (*TbEpsinR*) also localised to the endosomal compartment as well as the plasma membrane. RNAi knockdown suggested that *TbEpsinR* is required for endocytosis, as loss of this protein results in reduce intracellular pools of multiple surface antigens and cell growth defects in BSF (Gabernet-Castello et al., 2009). The data might indicate that the *TbVHS* is somehow associated with *TbRP2* as part of the trafficking pathway for ciliary proteins. As the base of the flagellum/cilium is undoubtedly a highly dynamic area it is possible that some of the cytosolic proteins identified in my study, reflect proteins that do not pass the ciliary gate and enter the flagellum (e.g. TCP-1-ε), whereas other proteins (e.g. *TbP25-α*, *TbBARTL1*) are transported into this compartment. Furthermore, as highly abundant proteins found in the flagellum (PFR proteins, αβ-tubulin), are not present in the *TbRP2*-proximity proteome, this suggests that the proteins identified by my study have an association with *TbRP2*.



**Figure 7.4 Working models of possible *TbRP2* interacting/nearby partners, which are identified by BioID-SILAC approach in *T. brucei*.** *TbRP2* is possibly in the centre of the complex with *TbBBP590*, *TbMARP*, *TbArl3L*, and *TbKMP11* at the basal body. *TbRP2* might involve in vesicle trafficking direct or indirect connection with *TbPK* and *TbVHS*. *TbMKS1* and *TbMKS6* were previously identified in the lab (Andre et al., 2014).

## Chapter 8 Future Work

In this study, all the selected putative proteins were suggesting that *TbRP2* is more likely to involve in general protein trafficking and act as a distinct module from the rest of known protein module at the base of the flagellum. In addition to a putative molecular client of *TbRP2* GAP activity (ARL3L) many of the tagged proteins identified in this study are consistent with *TbRP2* being located at a highly dynamic region of the cell involved in protein trafficking into the flagellum. In addition, there are proteins that could potentially be of importance as structural components of the basal body. However, many questions remain unanswered; therefore, the following experiments would be informative to expand upon the work described in this thesis:

- 1) Investigate other proteins identified for my BioID studies (see table 8.1). During the writing of my thesis, re-interrogation of these proteins was undertaken and I suggest that proteins such as the *ruvB*-like protein (*Tb927.4.2000*) should be prioritised for analysis. The *TbruvB*-like protein is homologous to human *ruvB*-like 2 isoform 1 (NP\_006657.1) and studies have shown that *RuvBL1* and another related *RuvB*-like protein *RuvBL2* along with other proteins form the R2TP complex (Boulon et al., 2008; Zhao et al., 2005). The R2TP complex is associated with many cellular processes (such as RNA polymerase II assembly) (Kakihara and Houry, 2012) but also interacts with Hsp90-associated proteins (R2TP-Hsp90 complex) and is implicated in ciliogenesis. Equally intriguing is the *Pih1/Nop17* protein (*Tb927.10.12860*) which shares ~ 25% identity with human protein *kintoun* (*Ktu/PF13*) isoform 2 (NP\_001077377.1). This protein belongs to the PIH1 family proteins, which are involved in pre-assembly of axonemal dynein arm. Mutations in *Ktu* cause primary ciliary dyskinesia (PCD) (due to an axonemal dynein arm defect) (Omran et al., 2008; Yamamoto et al., 2010). In mice, *Ktu/PF13* has been shown to act as a co-chaperone to interact with Hsp70, and is also involved in the cytoplasmic pre-assembly of dynein arms (Olcese et al., 2017).

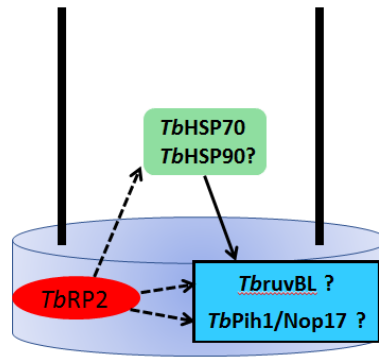


Figure 8.1 Proposed interaction/localisation model of ruvBL and pih1/Nop17 proteins to *TbRP2* and in HSP70/90-mediated axonemal dynein assembly, based on the studies in other models.

Protein ID	Accession Number	M/L Normalised Ratio (Cyto)	H/L Normalised Ratio (Cyto)	M/L Normalised Ratio (WC)	M/L Normalised Ratio (WC)	Kinetoplast-Specific Protein	<i>Homo Sapiens</i> (conserved)
<b>Table 3 : Proteins have NOT been selected for further studies, but with good SILAC enrichment</b>							
RNA editing associated helicase 2, mitochondrial RNA binding complex 1 subunit (REH2)	Tb 927.4.1500	0.78	4.30			No	KIAA1488 protein, partial (36%)
p197, a TAC component	Tb 927.10.15750	1.45	1.21	0.27	-0.26	Yes	No significant similarity found
Eukaryotic translation initiation factor 3 subunit F	Tb 927.3.1680			4.24	4.28	Yes	No significant similarity found
Hypothetical protein	Tb 927.9.16410 (previously known as : Tb 09.244.1650)			7.93	2.95	Yes	No putative conserved domains have been detected
Hypothetical protein	Tb 927.11.6440 (previously known as : Tb 11.02.4300)			2.13	2.31	Yes	No significant similarity found
Protein translation factor SUI1 homolog, putative	Tb 927.11.5840 (previously known as : Tb 11.02.3595)			2.22	2.22	No	translation factor sui1 homolog, partial [Homo sapiens] (76%)
Nucleosome Assembly protein-like protein	Tb 927.9.5730 (previously known as : Tb 09.160.4240)			2.17	2.18	No	nucleosome assembly protein 1-like 4b [Homo sapiens] (34%)
	Tb 927.8.2820			1.91	1.96	Yes	No significant similarity found
cystathionine beta-synthase, putative	Tb 11.02.5400b; Tb 11.02.5400			2.06	1.82	No	Chain A, Cystathionine-Beta Synthase: Reduced Vicinal Thiols
C-terminal motor kinesin, putative (TBKIFC1)	Tb 927.10.14890; Tb 11.v5.0636			1.42	1.77	No	Similar to kinesin family member C3, partial (45%)
hypothetical protein, conserved (pre-RNA processing PIH1/Nop17, putative)	Tb 927.10.12860			2.01	1.76	No	protein kintoun isoform 2 (25%)
zinc finger protein family member, putative (ZC3H41)	Tb 927.11.1980 (previously known as : Tb 11.46.0009)			1.86	1.63	Yes	
eukaryotic translation initiation factor 2 beta subunit, putative (EIF2B)	Tb 927.5.3120			1.71	1.56	No	EIF2S2 protein (32%)

ubiquitin-activating enzyme E1, putative (UBA2)	<i>Tb</i> 927.9.12650 (previously known as : <i>Tb</i> 09.211.3610)		1.43	1.44	No	ubiquitin activating enzyme E1 (34%)
pumilio RNA binding protein PUF1 (PUF1)	<i>Tb</i> 927.10.4430; <i>Tb</i> 10.v4.0033; <i>Tb</i> 11.v5.0767		2.03	1.40	No	pumilio homolog 2 isoform 3 (25%)
ruvB-like DNA helicase, putative,ATP-dependent DNA helicase, putative	<i>Tb</i> 927.4.2000		1.36	1.29	No	ruvB-like 2 isoform 1 (72%)
peptidase M20/M25/M40, putative	<i>Tb</i> 927.6.400		1.23	1.14	No	cytosolic non-specific dipeptidase isoform 1 (27%)
Calmodulin	<i>Tb</i> 927.11.13050; <i>Tb</i> 927.11.13040; <i>Tb</i> 927.11.13030; <i>Tb</i> 927.11.13020		1.12	1.14	No	Chain A, Crystal Structure Of The Plectin 1a Actin-binding Domain/n-terminal Domain Of Calmodulin Complex
NADH-dependent fumarate reductase (FRDg)	<i>Tb</i> 927.5.930; <i>Tb</i> 11.v5.0613		1.01	1.10	No	NADH-cytochrome b5 reductase 3 isoform 2 (28%)
Translation initiation factor eIF3	<i>Tb</i> 927.7.6090		0.76	1.04	Yes	No significant similarity found
Paraflagellar Rod component, putative (PFC14)	<i>Tb</i> 927.10.9570; <i>Tb</i> 11.v5.0579		1.19	0.55	Yes	
methionyl-tRNA synthetase, putative (MetRS)	<i>Tb</i> 927.10.1500		1.35	1.00	No	methionine-tRNA ligase, mitochondrial precursor (41%)

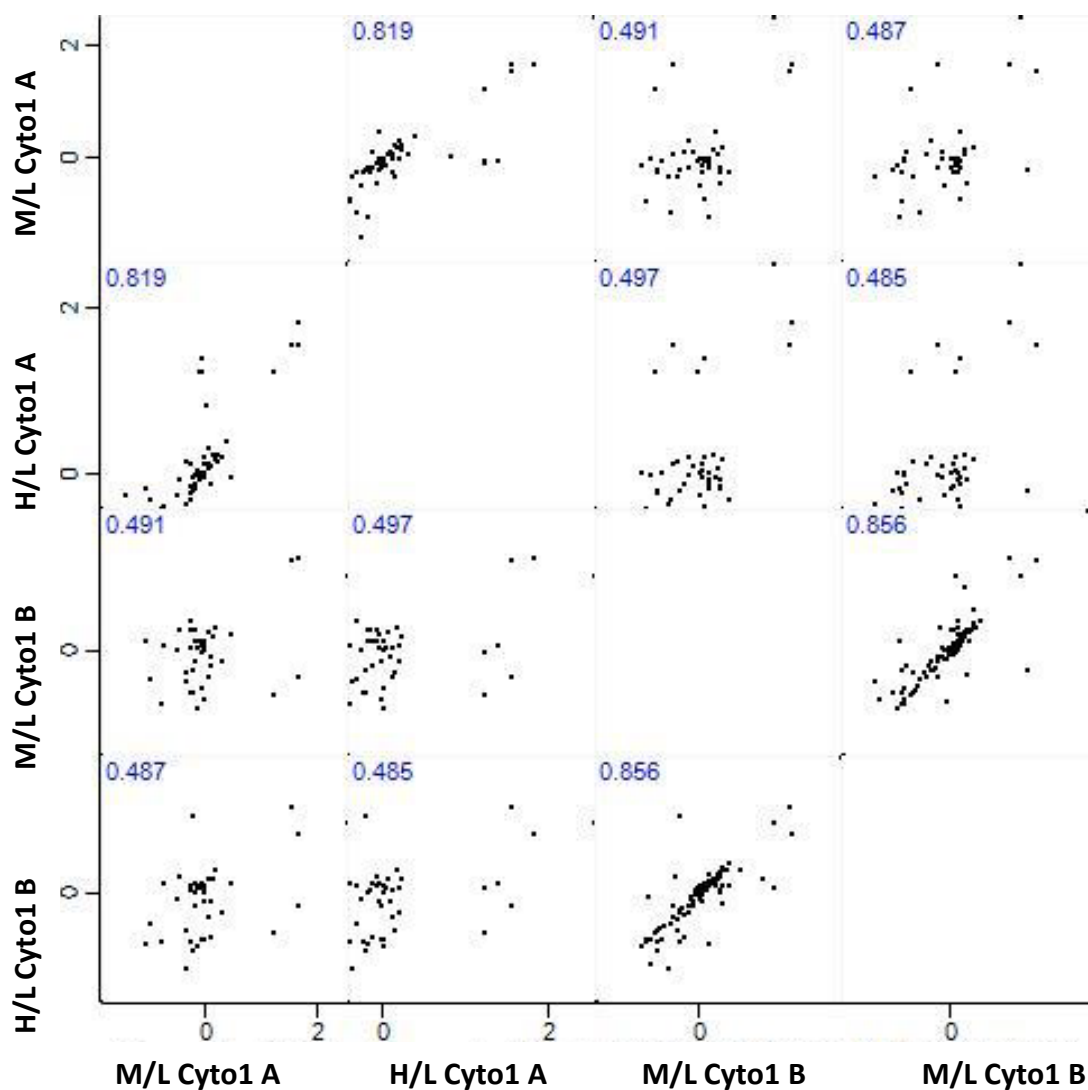
**Table 8.1** Selected proteins with good SILAC ratio (above 1). However, their functionality did not investigate in this study.



- 2) Loss of *TbRP2* leads to a short flagellum and axonemal defects (missing/mis-aligned central pair and disrupted outer doublet MTs) (Stephan et al., 2007). However, recent work from the laboratory of Philippe Bastin (Santi-Rocca et al., 2015) suggests that loss of *TbRP2* affects IFT. The effect of *TbRP2* RNAi on IFT and on the transport of flagellar membrane proteins needs to be explored further. For example, use of high throughput pPOT tagging techniques to tag candidate flagellum membrane proteins on a *TbRP2* RNAi background.
- 3) As mentioned in chapter 3, *T. brucei* cells expressing *TbRP2::*(T2)::myc::BirA\* and *TbRP2::*(T3)::myc::BirA\* exhibit abnormal morphology 24 hour after induction of protein expression (Fig 3.10 to Fig 3.13). As a result, these cell lines were not included in the BioID experiments. However, as expression of these truncated proteins results in a dominant negative phenotype, it would be interesting to examine whether the proteins identified by BioID/SILAC in these cells differed from those identified in this study. Similarly, expression of *TbRP2*<sup>R248H</sup> mutant protein (i.e. which will lack GAP activity) gives dominant negative effect when over-expressed on a wild type background. Therefore, it would be interesting to ask how the profile of proteins identified by BioID differed in cells where flagellum formation was compromised.
- 4) Reciprocal tagging analysis, for example by investigating putative interacting/nearby proteins of *TbHSP70::*myc::BirA\* and *KMP-11::*myc::BirA\* to check if *TbPR2* or any other identified proteins would appear in their protein profile.
- 5) To understand and expand our knowledge of the protein complex at the base of the flagellum or how their might involve in flagellum assembly. We could use a systematic BioID analysis to use other TZ or basal body proteins as baits.

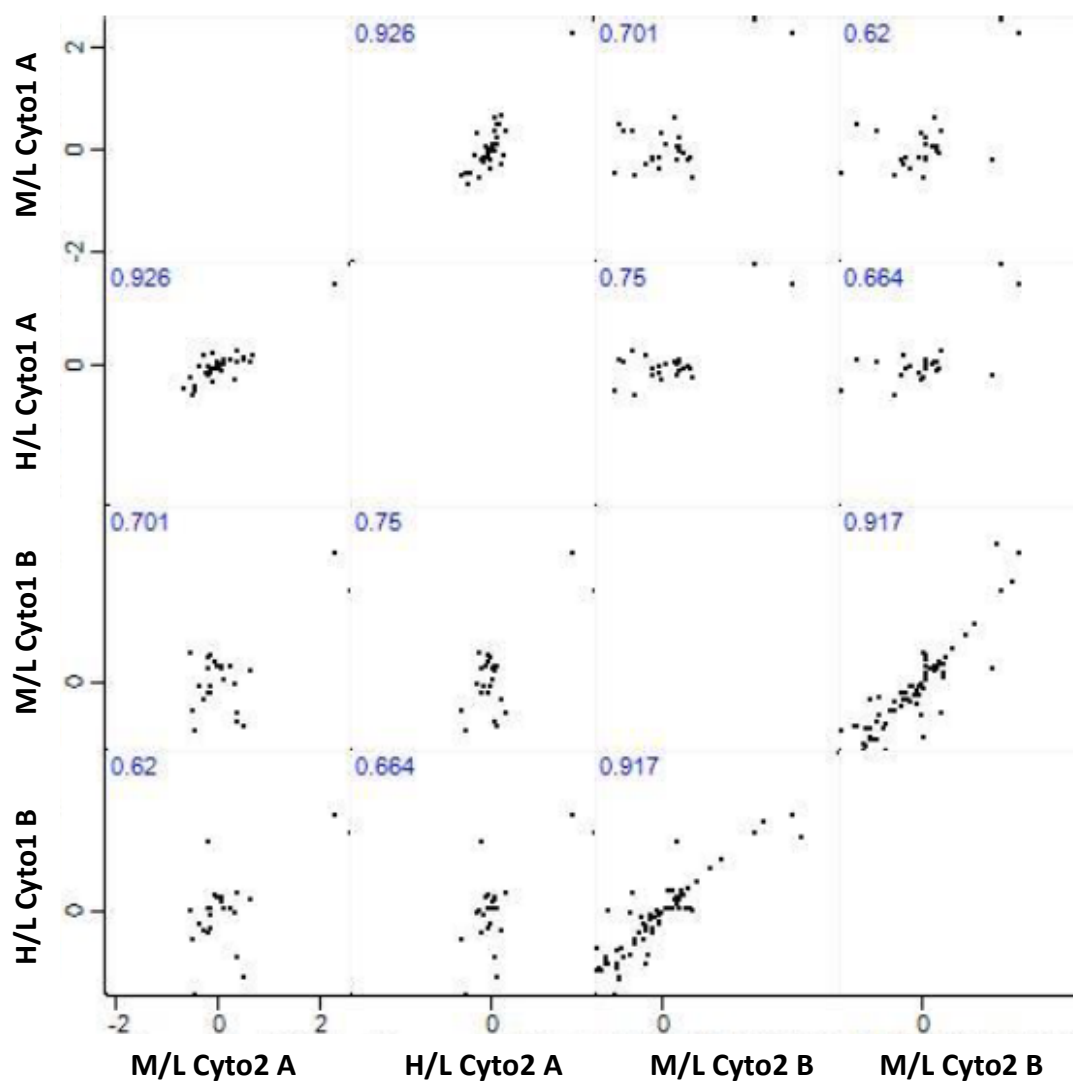
## Chapter 9 Appendices

### Appendix I



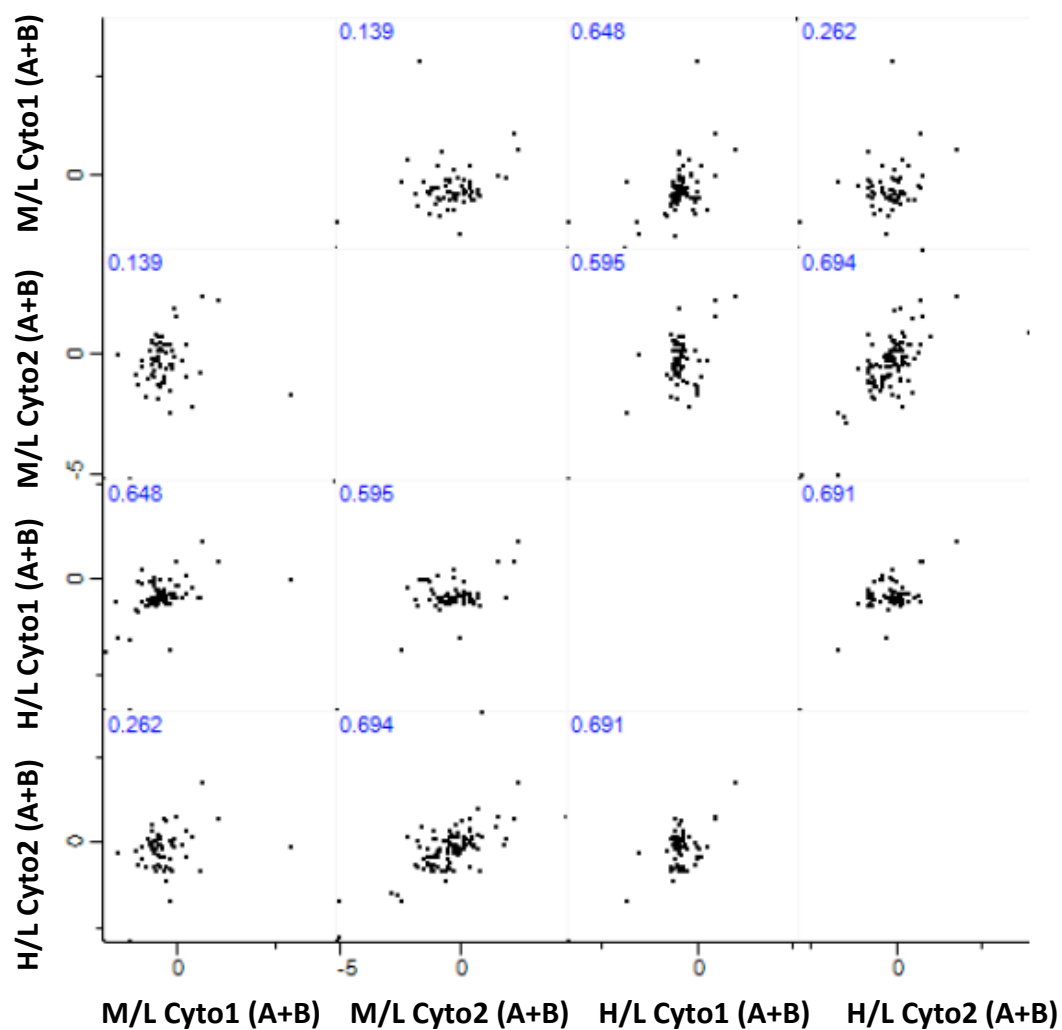
	Cyto1 A (M/L)	Cyto1 A (H/L)	Cyto1 B (M/L)	Cyto1 B (H/L)
Cyto1 A (M/L)		0.819	0.491	0.487
Cyto1 A (H/L)	0.819		0.497	0.485
Cyto1 B (M/L)	0.491	0.497		0.856
Cyto1 B (H/L)	0.487	0.485	0.856	

**Fig 9.1 Correlation of technical replicate experiments of Expt cyto-1.** Pair wise Pearson correlations of the Log2 M/L (*TbRP2*<sup>Δ134-463</sup>/PTP) ratio and H/L (*TbRP2* FL/PTP) ratio of two individual experiments (A and B) of the same sample set. Table shows the Pearson's correlation value, which were generated by *Perseus* software.



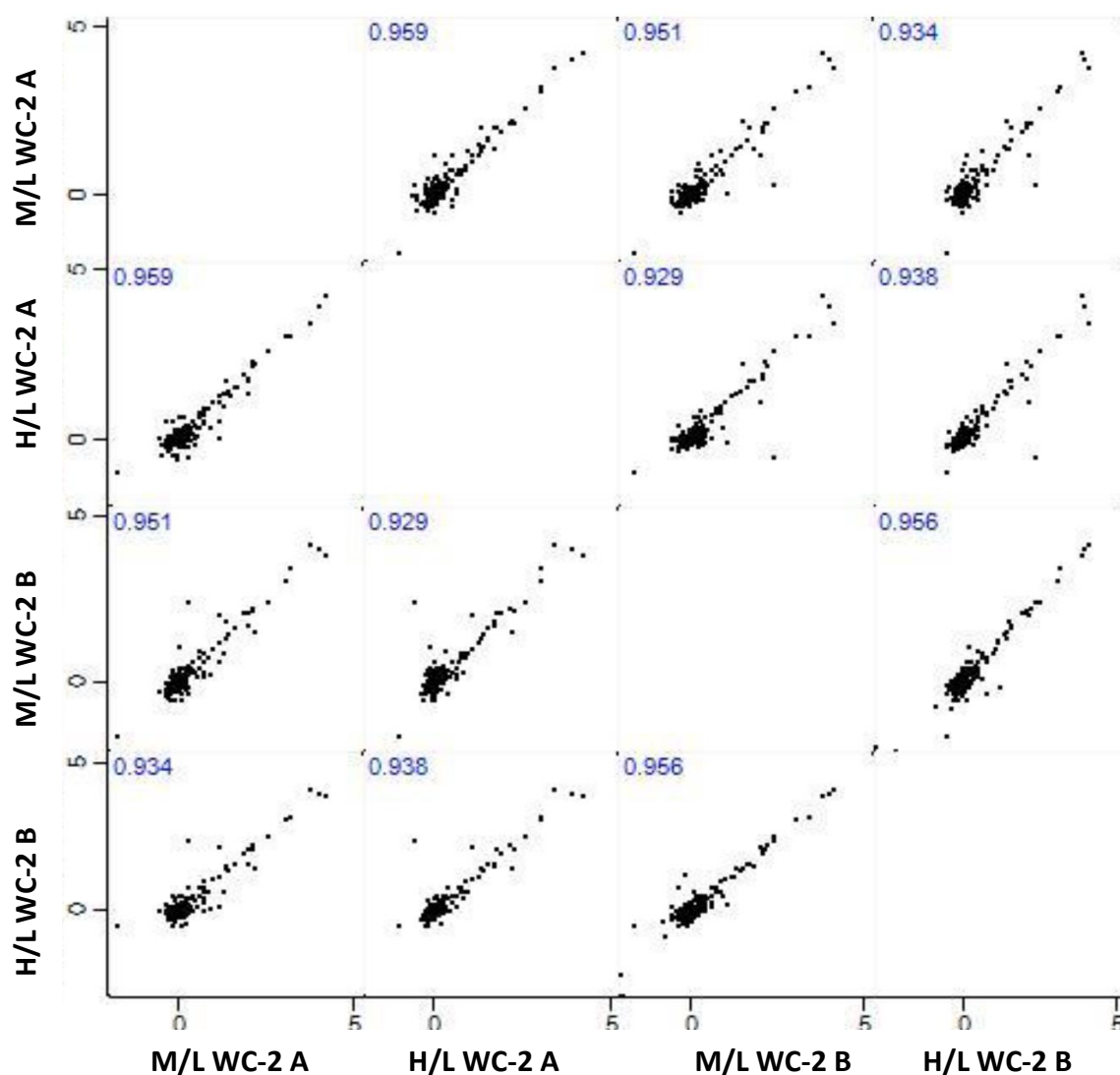
	Cyto2 A (M/L)	Cyto2 A (H/L)	Cyto2 B (M/L)	Cyto2 B (H/L)
Cyto2 A (M/L)		0.926	0.701	0.660
Cyto2 A (H/L)	0.620		0.750	0.664
Cyto2 B (M/L)	0.701	0.750		0.917
Cyto2 B (H/L)	0.926	0.664	0.917	

**Fig 9.2 Correlation of technical replicate experiments of Expt cyto-2.** Pair wise Pearson correlations of the Log2 M/L (*TbRP2*<sup>Δ134-463</sup>/PTP) ratio and H/L (*TbRP2* FL/PTP) ratio of two individual experiments (A and B) of the same sample set. Table shows the Pearson's correlation value, which were generated by *Perseus* software.



	Cyto1 A+B (M/L)	Cyto1 A+B (H/L)	Cyto2 A+B (M/L)	Cyto2 A+B (H/L)
Cyto1 A+B (M/L)		0.648	0.139	0.262
Cyto1 A+B (H/L)	0.648		0.595	0.691
Cyto2 A+B (M/L)	0.139	0.595		0.694
Cyto2 A+B (H/L)	0.262	0.691	0.694	

**Fig 9.3 Correlation of biological replicate experiments of Expt cyto-1 and Expt cyto-2.** Pair wise Pearson correlations of Log2 M/L (*TbRP2*<sup>Δ134-463</sup>/PTP) ratio and H/L (*TbRP2* FL/PTP) ratio of combined data (e.g. M/L of Cyto1 A+B and M/L of Cyto2 A+B as a set of biological replicate). Table shows the Pearson's correlation value, which were generated by *Perseus* software.



	WC-2 A (M/L)	WC-2 A (H/L)	WC-2 B (M/L)	WC-2 B (H/L)
WC-2 A (M/L)		0.959	0.951	0.934
WC-2 A (H/L)	0.959		0.929	0.938
WC-2 B (M/L)	0.951	0.929		0.956
WC-2 B (H/L)	0.934	0.938	0.956	

**Fig 9.4 Correlation of technical replicate experiments of Expt WC-2 A and B.** Pair-wise Pearson correlations of the Log2 M/L (*TbRP2* FL/PTP) ratio and H/L (*TbRP2* FL/PTP) ratio from Expt WC-2 A and B (label swap). Table shows the Pearson's correlation value, which were generated by *Perseus* software.

## Appendix II

Appendix II contains multiple amino acid sequence alignments of some of the selected putative *TbRP2*-interacting proteins in either (1) *T. brucei* only; or (2) trypanosomatids only, or (3) compared with other organisms (Table 4.2). Amino acid identity (*asterisk*) and similarity (:) between sequences are indicated.

Grp170 -----MRVLAN-----EIRRM-----LVAFLLTAA  
HSP70.a -----MSMLCL-----AQWAL-----LLVLCLVGC  
HSP70.b MHCKVFSFRLTAGLLQRLPMPMPPPRPGVSVNMQHAVEIEAKRRVELDEATRARYVVVK  
HSP110 -----  
mtHSP70A -----M-----LARRVCA-PMCLASAPFAR  
mtHSP70B -----M-----LARRVCA-PMCLASAPFAR  
mtHSP70C -----M-----LARRVCA-PMCLASAPFAR  
Hsp70.c -----  
Bip/Grp78B -----MSRMWL-----TTAAVFL-TVTVAAVSAAP  
Bip/Grp78A -----MSRMWL-----TTAAVFL-TVTVAAVSAAP  
HSP70 -----  
HSP70.4 -----

Grp170 FGYDVTSANVIGVDFGSDYIE---VAGPIN--GVNVDIILNEQSHRKT DNYIGFRN---  
HSP70.a CCTVSGGSEVLAVDIGADWAKGATRVIGGST--APRASIVLNDQTNRKSPQCI AFRI VPN  
HSP70.b EETRASGDRVIGIDLGT TNSC-----ISYIDKKTNRPKIIPSPTG SWVFPTAITFDKS--  
HSP110 -----MSVFGVDFGNLNTT-----VAITR--HGGVDIVTNEVSKRETTT IASFVDN--  
mtHSP70A WQSSKVTGDVIGIDLGT TYSV-----VAVME--GDRPRVLENT EGFRTTPSVVAFKQ--  
mtHSP70B WQSSKVTGDVIGIDLGT TYSV-----VAVME--GDRPRVLENT EGFRTTPSVVAFKQ--  
mtHSP70C WQSSKVTGDVIGIDLGT TYSV-----VAVME--GDRPRVLENT EGFRTTPSVVAFKQ--  
Hsp70.c ----MTYEGAIGIDLGT TYSV-----VG VWQ--NERVEIIANDQGNRTTPSYVAFVNN--  
Bip/Grp78B ESGGKVEAPCVGIDLGT TYSV-----VG VWQ--KGDVHII PNEMGNRITPSVVAFTDT--  
Bip/Grp78A ESGGKVEAPCVGIDLGT TYSV-----VG VWQ--KGDVHII PNEMGNRITPSVVAFTDT--  
HSP70 ----MTYEGAIGIDLGT TYSV-----VG VWQ--NERVEIIANDQGNRTTPSYVAFVNN--  
HSP70.4 ----MPAPAIGIDLGT TYSV-----VG VFK--NDQVEIVANDQGNRTTPSYVSFSET--  
. . : \* : \*

Grp170 -----GERSIGAQA KSLAARFPTNMIAMINHLVGITYNSSDFAN-FKKLQCEFDPH P  
HSP70.a AGNDTLRSVERLFAEEARSLEPRFPQQSICGPSLLAGLIVSKEISAG-QKHHEQTGNQRS  
HSP70.b -----HKVRLYGEEARACVRTSASATLC SGKRLIGRGV GELGRVQSQLHKT NMVTLN-  
HSP110 -----ERFIGEPGLDRYVRNSSNTIFLLKRFIGMYMDDPSLESEERRFLTCAIKGD-  
mtHSP70A -----EKLVLGLAAKRQAITNPQSTFFAVKRLIGRRFDEHIQHDIKNVPYKIIRS-  
mtHSP70B -----EKLVLGLAAKRQAITNPQSTFFAVKRLIGRRFDEHIQHDIKNVPYKIIRS-  
mtHSP70C -----EKLVLGLAAKRQAITNPQSTFFAVKRLIGRRFDEHIQHDIKNVPYKIIRS-  
Hsp70.c -----EVLVGDAAKNHAARGSNVIFDAKRLIGRKFSDSVQSDMKHWPFPKVEEG-  
Bip/Grp78B -----ERLIGDGAKNQLPQNPHNTIYTIKRLIGRKYTDAAVQADKKLLSYEVIAD-  
Bip/Grp78A -----ERLIGDGAKNQLPQNPHNTIYTIKRLIGRKYTDAAVQADKKLLSYEVIAD-  
HSP70 -----ERLIGDGAKNQVAMNPNTNTVFD AKRLIGRKFSDSVQSDMKHWPFPKVTK-  
HSP70.4 -----ERLVGDAAKNQVAMNPNTNTVFD AKRIIGRKYDDPDLQADMKHWPFPKVTK-  
. . . : \*

Grp170 EERGTVGFRFE-----DNNDTYTAEEIYAMMLNYC  
HSP70.a EREGVISFSDTDRFTYVVVPQIRRKSAVVRITPGGSSEGT TTAPIETVEELIGMILGHM  
HSP70.b -ERGEVAVEI-----MGRTYTVTHIIAMFLRYL  
HSP110 -DKGRLMFGV-----NY-----CGELTYFYPEQVLAMMLQRL  
mtHSP70A -NNGDAWVQ-----DGNGKQYSPSQVGAFVLEKM  
mtHSP70B -NNGDAWVQ-----DGNGKQYSPSQVGAFVLEKM  
mtHSP70C -NNGDAWVQ-----DGNGKQYSPSQVGAFVLEKM  
Hsp70.c -EKGGAVMRV-----EH-----LGEGMLLQPEQISARVLAYL  
Bip/Grp78B -RDGKPKVQV-----MV-----GGKKKQFTPEEISAMVLQKM  
Bip/Grp78A -RDGKPKVQV-----MV-----GGKKKQFTPEEISAMVLQKM  
HSP70 -GDDKPVIVQV-----QF-----RGETKTFNPPEEISSMVL LKM  
HSP70.4 --EGKPVVEV-----EY-----QGERRTFPEEISAMVLQKM  
. . : . \*

Grp170 RSI-----SEKA-GVPNPQNVVITLPHFSSLRGRRQTILEAARLVHINTLGLLHSTTATA  
HSP70.a KRS-----AERSLDGAPVRHLVLVVP TSSSLAYRQAMVDAAAVGLRTIRLVHGSAAAA  
HSP70.b KKE-----AEKF-LKEPVNAVVSVPAFFTPQKQVATEDAALAAGFDVLEVIDEPSAAC  
HSP110 RGYVNLASLSDSK-VTVDSRECVLTVPCYYTAEQRKLLMQACEIAGLNCLSLVNDTTAAG  
mtHSP70A KET-----AENF-LGRKVSNAVVTCPAYFNDAQRQATKDAGTIAGLNVRVNEPTAAA  
mtHSP70B KET-----AENF-LGRKVSNAVVTCPAYFNDAQRQATKDAGTIAGLNVRVNEPTAAA  
mtHSP70C KET-----AENF-LGRKVSNAVVTCPAYFNDAQRQATKDAGTIAGLNVRVNEPTAAA  
Hsp70.c KSC-----AESY-LGKQVAKAVVTVPAYFNDSQRQATKDAGTIAGLEVLRINEPTAAA  
Bip/Grp78B KEI-----AETY-LGEKVKNNAVVTVPAYFNDAQRQSTKDAGTIAGLNVVRIINEPTAAA  
Bip/Grp78A KEI-----AETY-LGEKVKNNAVVTVPAYFNDAQRQSTKDAGTIAGLNVVRIINEPTAAA  
HSP70 KEV-----AESY-LGKQVAKAVVTVPAYFNDSQRQATKDAGTIAGLEVLRINEPTAAA  
HSP70.4 KEI-----AESY-LGEKVSNAVVTVPAYFNDSQRQATKDAGSIAGLEVLRINEPTAAA  
: : : \* \* . : : \* . : : : : \*

Grp170 LYYGV-----RRR-GF---GNRTVHLLVYDIGSTHTEVG VYKFSPPVQEQGKRKTKNVESF  
HSP70.a TQLAHLNTETLFR-GHPSNTTERKYAMIYDMGSSKTEVAVFRFTPATAR-----DDF

HSP70.b LAHTVLQPSNASSREHLSGSKRIVRSLVFDLGGGTLDCAVMEND-----RRR  
HSP110 IDYGI FRGSSSLGETE---D---KGQVVGILDMGYGTTVFVAVACF-----WR  
mtHSP70A LAYGLDKTK-----DSLIAVYDLGGGTFDISVLEI-----AG  
mtHSP70B LAYGLDKTK-----DSLIAVYDLGGGTFDISVLEI-----AG  
mtHSP70C LAYGLDKTK-----DSLIAVYDLGGGTFDISVLEI-----AG  
Hsp70.c IAYGLDKAD-----E---G---KERNVLVDFDGGGTFDVSISV-----SG  
Bip/Grp78B IAYGLNKGK-----EKNILVFDLGGGTFDVSLLTI-----DE  
Bip/Grp78A IAYGLNKGK-----EKNILVFDLGGGTFDVSLLTI-----DE  
HSP70 IAYGLDKAD-----E---G---KERNVLIFDLGGGTFDVTLLTI-----DG  
HSP70.4 IAYGMDRSS-----E---G---AMKTVLIFDLGGGTFDVTLLNI-----DG  
: \* : \*

Grp170 GTLTTMGIVSDATLGGRAIDSCIAKGIEAEAIKMKISP---VLGGS--TVSQRKAQFSL  
HSP70.a GTVTLVASATNHTLGGRSFDRCLARYVERNLFPAAKPTPVPVLDKRPVTATTRRAVVS  
HSP70.b GTFTLVATHGDPPLGGNDWDAVLSQHFSQDQFER-KWRVPLE---DAEGNVGQGVATYRQL  
HSP110 GHLKLSRTFDRHLGTRDIDYKLFYMAEEVKK-KYHVDVK---ENK-----RASLRL  
mtHSP70A GVFEVKATNGDTHLGGEDFDLCLSDHILEEFRK-TSGIDLS---KER-----MALQRI  
mtHSP70B GVFEVKATNGDTHLGGEDFDLCLSDHILEEFRK-TSGIDLS---KER-----MALQRI  
mtHSP70C GVFEVKATNGDTHLGGEDFDLCLSDHILEEFRK-TSGIDLS---KER-----MALQRI  
Hsp70.c GVFEVKATNGDTHLGGEDVDAALLEHALADIRNRYGIEQGS---LSQ-----KMLSKL  
Bip/Grp78B GFFEIVATNGDTHLGGEDFDNNMMRHVFDMLKK-KKNVDIS---KDQ-----KALARL  
Bip/Grp78A GFFEIVATNGDTHLGGEDFDNNMMRHVFDMLKK-KKNVDIS---KDQ-----KALARL  
HSP70 GIFEVKATNGDTHLGGEDFDNRLVAHFTEEFKRKNKGKDL---SNL-----RALRRL  
HSP70.4 GLFEVRATAGDTHLGGEDFDSRLVDYFATEFRT-RTGKDLR---GNA-----RAMRRL  
\* . . : \*\* . \* :

Grp170 FRAAKRAREVLSVNSKTPVTEGIAPD-----RDFSTEVS RKEFETNCSELLKRFPRV  
HSP70.a LRAVNAARERLSVNQNVFPVPGVRED---G---GDFIANISRAQFEEACGELFNEAVRL  
HSP70.b LLEAEKAKIHFTHSTEPYGYGNRAHFHSEKLRDIVPLEATLTLEEYIELTRPLVRVCEC  
HSP110 LQACERLRYLLSGNQVAQLNVENLMDV-----DVNIPSPFPRSTLEELSVCLVERFKAV  
mtHSP70A REAAEKAKCELSTTMETEVNLPFITANQDGA---QHVQMMVSRSKFESLADKLVQRS LGP  
mtHSP70B REAAEKAKCELSTTMETEVNLPFITANQDGA---QHVQMMVSRSKFESLADKLVQRS LGP  
mtHSP70C REAAEKAKCELSTTMETEVNLPFITANQDGA---QHVQMMVSRSKFESLADKLVQRS LGP  
Hsp70.c RSRCEEVKRVLSHSTVGEI ALDGLLPD---G---EEYVLKLTARLEELCTKIFARCLSV  
Bip/Grp78B RKACEAAKRQLSSHPEARVEVDSL-T-E---G---FDFSEKITRAKFEELNMDLFGKTLVP  
Bip/Grp78A RKACEAAKRQLSSHPEARVEVDSL-T-E---G---FDFSEKITRAKFEELNMDLFGKTLVP  
HSP70 RTACERAKRTLSSAAQATIEIDALF-E---N---IDFQATITRARFEELCGDLFRGTLQP  
HSP70.4 RTACERVKRTLSSASTNIEIDALY-E---G---FDFFSKITRARFEEMCRDQFERCLEP  
: : : :

Grp170 AQEAVTKSSL---TLKDIDAFEMMGASRTPKIIISDLSTFWG-KEV-NRTLNSDEAAAMG  
HSP70.a RDHAITQTNGTVRSLNELVRLELIGGATRM PKLQERLSEGYG-KPA-DRTLNSDEAVVSG  
HSP70.b LNKLF DHTSI---RPADIDNVLLVGAMTRDPPIRHLLTEYFGRHVESEASCPADYAVAIG  
HSP110 IKKGFEESGL---SPDQFHSIEMIGGGSRI PMFKSAAEELLG-RAP-NFTLNASETAARG  
mtHSP70A CKQCIKDAAV---DLKEISEVVLVGGMTRMPKVVEAVKQFFG-REP-FRGVNPDEAVALG  
mtHSP70B CKQCIKDAAV---DLKEISEVVLVGGMTRMPKVVEAVKQFFG-REP-FRGVNPDEAVALG  
mtHSP70C CKQCIKDAAV---DLKEISEVVLVGGMTRMPKVVEAVKQFFG-REP-FRGVNPDEAVALG  
Hsp70.c VQRALKDASM---KVEDIEDVVLVGGSSRI PAVQAQLRELFRGKQL-CSSVHPDEAVAYG  
Bip/Grp78B VQRVLEDAKL---KKSIDIHEIVLVGGSTRVPKVQQLISDFFGGKEL-NRGINPDEAVAYG  
Bip/Grp78A VQRVLEDAKL---KKSIDIHEIVLVGGSTRVPKVQQLISDFFGGKEL-NRGINPDEAVAYG  
HSP70 VERVLDQAKM---DKRAVHDVVLVGGSTRIPKVMQLVSDFFGGKEL-NKSINPDEAVAYG  
HSP70.4 VRKVLKDAEV---DASAVDDVVLVGGSTRIPRVQQLVQNFENGKEP-NRSINPDEAVAYG  
. . . : . . : \* . : \* . : \* . : \* . : \*

Grp170 AAYYALRLSPHYRPH---SFRIIERVPYTF LF-----AVSPEVNKSSNSKR  
HSP70.a AALMIHD-----TLS-----RIRVMESLTNDIYF-----TASPPIKESNETKP  
HSP70.b AAVRGAMLQGGFDDLLSNTRFVTGTA---QALKQGG-----TASPPIKESNETKP  
HSP110 AAITAAYVSPKFK---VREFVVS DIPTYPIKLGYYMENASAVSHVPFLPDINKVVSQVQ  
mtHSP70A AATLGGVLRGD-----VKGLVLLDV---TPLSLGIETLGGVFTRMIP-----TASPPIKESNETKP  
mtHSP70B AATLGGVLRGD-----VKGLVLLDV---TPLSLGIETLGGVFTRMIP-----TASPPIKESNETKP  
mtHSP70C AATLGGVLRGD-----VKGLVLLDV---TPLSLGIETLGGVFTRMIP-----TASPPIKESNETKP  
Hsp70.c AAVQAHLVSGGYGESSRTAGIVLLDV---VPLSIGVEVDGKFDVIR-----TASPPIKESNETKP  
Bip/Grp78B AAVQAAVLTGE---SEVGGRVVLVDV---IPLSLGIETVGGVMTKLIE-----TASPPIKESNETKP  
Bip/Grp78A AAVQAAVLTGE---SEVGGRVVLVDV---IPLSLGIETVGGVMTKLIE-----TASPPIKESNETKP  
HSP70 AAVQAFILTGGK---SKQTEGLLLLDV---APLTLGIETAGGVMTALIK-----TASPPIKESNETKP  
HSP70.4 AAVQAHIVSGGK---SKQTKDLLLDV---TPLSLGVETAGGVMSVLIP-----TASPPIKESNETKP  
\*\* . :

Grp170 M-----LANNPIIGSRWISITVNR---TDDFSIQLFDS DARIGN-VNITGVKGALERLGF  
HSP70.a HRNLLFAKRNTTVPAARSLIFPNR---TADFTLT LH DGNRGYRSRVLSVSGVSGSMNAARE  
HSP70.b -----FLRR-----CCNRIGS  
HSP110 --TDDHYPKVLEITIKRPGGFKLY-----AFYDSEHPKV KAYLPR-----KDFVIGE



```

mtHSP70A -----KNTTIPTKKSQTFSTAADNQTQVGIKVFQGEREM-----AS----DNQMMGQ
mtHSP70B -----KNTTIPTKKSQTFSTAADNQTQVGIKVFQGEREM-----AS----DNQMMGQ
mtHSP70C -----KNTTIPTKKSQTFSTAADNQTQVGIKVFQGEREM-----AS----DNQMMGQ
Hsp70.c -----RNTTIPYLATKEYSTVDDNQSEVEIQVFEGGERPL-----TR----HNHRLGS
Bip/Grp78B -----RNTQIPTKKSQVFSTHADNQPGVLIQVYEGERQL-----TK----DNRLLGK
Bip/Grp78A -----RNTQIPTKKSQVFSTHADNQPGVLIQVYEGERQL-----TK----DNRLLGK
HSP70 -----RNTTIPTKKSQIFSTYSNQPVGHIQVFEGERTM-----TK----DCHLLGT
HSP70.4 -----RNTSVPAQKSQTFSTNADNQSRVEIKVYEGERPL-----VS----QCQCLGT

```

```

Grp170 FQPKLNHTNNSHI-----IRIQVRFNESGLLEVEEAGVFY-----RYAVNVS--
HSP70.a KEKEMSTERANKVTKTSVVLRLQVEVVVVLSRSGLPYVAGSYVHA-----RYAEQVTVL
HSP70.b LV--SSSVNP-----NAIGQ-----RWRGRAKGLSDEEIANYAKELVEFEACDRLL
HSP110 WE--IGTQR-----KDSNATEVRVRVRLPLNGLVSV-----ESAVSVEVY
mtHSP70A FD--LVGIPP-----APRGVPQIEVTFDIDANGICHV-----TAK-----
mtHSP70B FD--LVGIPP-----APRGVPQIEVTFDIDANGICHV-----TAK-----
mtHSP70C FD--LVGIPP-----APRGVPQIEVTFDIDANGICHV-----TAK-----
Hsp70.c FV--LDGITP-----AKHGEPTITVTFSDADGILTV-----TAA-----
Bip/Grp78B FE--LSGIPP-----AARGVPQIEVTFDIDENSILQV-----SAM-----
Bip/Grp78A FE--LSGIPP-----AARGVPQIEVTFDIDENSILQV-----SAM-----
HSP70 FD--LSGIPP-----APRGVPQIEVTFDLANGILSV-----SAE-----
HSP70.4 FT--LTDIPP-----APRGKPRITVSFDVNVGDILV-----TAV-----

```

...

```

Grp170 -----SSTKTTGVQADSNS-----SNESG--YVEMK-----R--
HSP70.a -----PSVKKTGDNETTAQKDENNPSQNETD--TTSTISP GREKRS G
HSP70.b LERAENDANFVMRRVTADSSKRQGMQE-----
HSP110 EVEEPAD-----AEEGGKGETANEEGEQPAEKKPMVKKQKQRRVELSVTPRL-----
mtHSP70A -----DKATGKTQNI-----TI--T-----
mtHSP70B -----DKATGKTQNI-----TI--T-----
mtHSP70C -----DKATGKTQNI-----TI--T-----
Hsp70.c -----EELGSVTKTL-----VV--E-----
Bip/Grp78B -----DKSSGKKEEI-----TI--TN-----
Bip/Grp78A -----DKSSGKKEEI-----TI--TN-----
HSP70 -----EKGTGKRNI-----VI--TN-----
HSP70.4 -----EETAGKTQAI-----TI--SN-----

```

```

Grp170 -----KSARLS---ARVTFTNPEPLSYELL-----NNSRVKIFQLLEK-
HSP70.a GSPSAANSNSAKMQNSRADEAKENETPTGDEILEVNERDAGTGGKNNNAKVRHFALRFP
HSP70.b -----KRVRLSEQLK-----FWQYMVHN
HSP110 -----DVIGLTGQEI V-----EFQKKETE
mtHSP70A -----AHGGLTKEQIE-----NMIRDSEM
mtHSP70B -----AHGGLTKEQIE-----NMIRDSEM
mtHSP70C -----AHGGLTKEQIE-----NMIRDSEM
Hsp70.c -----NSERLTSEEVQ-----KMIEVAQK
Bip/Grp78B -----DKGRLSEEEIE-----RMVREAAE
Bip/Grp78A -----DKGRLSEEEIE-----RMVREAAE
HSP70 -----DKGRLSKADIE-----RMVSDAAK
HSP70.4 -----DKGRLSREQID-----KMVAEAEK

```

::

```

Grp170 -----ERKKHEAATAKNNLETY-----LFWAKTEGI-LENTTALE
HSP70.a NNTPAPSSTSQGGVNMNKEEALAARNRLRALQRLDDERLRRSGLLNDVESL-LLHYKSLD
HSP70.b F-----HDHEDELLRTVRELEQALDELEGLAEDNTSGLTTAGTVDFSSVTPVNHCEEEE
HSP110 M-----ND-RDALI---TKTRDSKNELESYILDNRPRIADGGIL-----
mtHSP70A H-----AE-ADRVK---RELVEVRNNAETQANTAERQLTEWKYV-----
mtHSP70B H-----AE-ADRVK---RELVEVRNNAETQANTAERQLTEWKYV-----
mtHSP70C H-----AE-ADRVK---RELVEVRNNAETQANTAERQLTEWKYV-----
Hsp70.c F-----AL-TDATA---LARMEATERLTQWFDRLAEMETVP-----
Bip/Grp78B F-----ED-EDRKV---RERVDARNSLESVAYSLRNQVNDKDKL-----
Bip/Grp78A F-----ED-EDRKV---RERVDARNSLESVAYSLRNQVNDKDKL-----
HSP70 Y-----EA-EDKAQ---RERIDAKNGLENYAFSMKNTINDPN-V-----
HSP70.4 F-----AE-EDRAN---AEKIEARNSVENYTFSLRSTLSDPD-V-----

```

:

```

Grp170 PYTP----EEVNSLKNALAEAEQWLEDGDGSNE-SCSKEEYENKLATLKELVKNSGKN--
HSP70.a AWSAQQSDDNSNDWRSVVKDVSRRWFEEVGGDVN-VTELQ---KQYQRLKDLKLGE-----
HSP70.b RDCSSVSAASRSAQLRT-AHGDGKLKERTQDEEG--EKPKGRKIMRR-----
HSP110 --CEYVTKEQQAQFIQLANEYENWLYEDGADA-ELNVYQERVKTLRAIADAASDRRRNFE
mtHSP70A -----TDAEKENVRTLLA-ELRKVMENPNVTKD--ELSASTDKLQKAVM-----
mtHSP70B -----TDAEKENVRTLLA-ELRKVMENPNVTKD--ELSASTDKLQKAVM-----

```

```

mtHSP70C      -----TDAEKENVRTLLA-ELRKVMENPNVTKD--ELSASTDKLQKAVM-----
Hsp70.c      ---QPYSEKL-QKRIAFPHGKEWVGTLHTYTDAAASIEAKVAKIERLAK-----
Bip/Grp78B    --GGKLDPNDKAAVETAVAEAIRFLDENPNAEKE--EYKTALETTLQSVTN-----
Bip/Grp78A    --GGKLDPNDKAAVETAVAEAIRFLDENPNAEKE--EYKTALETTLQSVTN-----
HSP70         --AGKLDDADKNAVTTAVEEALRWLNDNQEASLD--EYNHRQKELEGVCA-----
HSP70.4       --QQNISQEDQQKIQTIVNAVVNWLDENRDATKE--EYDAKNKEIEQVAH-----

Grp170        -----KNTTNTTSGSD-----DAVKG-----
HSP70.a       -----
HSP70.b       --AVPLPRASAEAQ-----ELVEA-----GHPALRGADVSMTE-----
HSP110        DVEFELPTFKQEVNKAKNATALAAIGKEAHITEEELRGAAAKCDEALAWAENEMEKYRKQH
mtHSP70A      -----
mtHSP70B      -----
mtHSP70C      -----
Hsp70.c       -----
Bip/Grp78B    -----
Bip/Grp78A    -----
HSP70         -----
HSP70.4       -----

Grp170        -----DL-----
HSP70.a       -----
HSP70.b       ----ST-RSAFFEAQ----VEERAWRE-PPTP-PG-EHGSW-QEVKRAVDAGEPVGSP
HSP110        KSESPVLTCTSLREKQKEVAEAVRAVVKRPPPKPKPEATPEA-KDGETAADGNVP----
mtHSP70A      ----ECGRTE---Y-Q-----QAAAAAN-----S-GSSGS---SSTEGQGEQQQQQAS
mtHSP70B      ----ECGRTE---Y-Q-----QAAAAAN-----S-GSSGS---SSTEGQGEQQQQQAS
mtHSP70C      ----ECGRTE---Y-Q-----QAAAAAN-----S-GSSGS---SSTEGQGEQQQQQAS
Hsp70.c       ----RALKSA---R-----RE-GKDGWAPGNEDNGSGDDNDGDDN
Bip/Grp78B    ----PIIQKT---Y-Q-----SAGGGDKPQPM-DD-L-----
Bip/Grp78A    ----PIIQKT---Y-Q-----SAGGGDKPQPM-DD-L-----
HSP70         ----PILSKM---Y-Q-----GMGGGDAAGGM-PG-GMPGGMPGGMPGGMGGMGGAAA
HSP70.4       ----PILSAY---Y-V-----KRAEQAPPAP-PS-GE-----GEGN-----

Grp170        -----
HSP70.a       -----
HSP70.b       GLQELQRPMTHEEMLQVLN----NIAPIDDPVSEEHARKRDHSIDMRTMTIVEGAVDMV
HSP110        -----PEVSQAG----DGEKPPSDDQLD-----
mtHSP70A      -----GEKKE-----
mtHSP70B      -----GEKKE-----
mtHSP70C      -----GEKKE-----
Hsp70.c       -----SDEDELQGRGVTEGSGRPPIRKRDRIEA-----INANTE-----
Bip/Grp78B    -----
Bip/Grp78A    -----
HSP70         -----SSGPKVEEVD-----
HSP70.4       -----APVPDDVD-----

Grp170        -----
HSP70.a       -----
HSP70.b       ALQELLEEAAKRAEELQRAQKKGEKQLVADSSAKLFAMD
HSP110        -----
mtHSP70A      -----
mtHSP70B      -----
mtHSP70C      -----
Hsp70.c       -----
Bip/Grp78B    -----
Bip/Grp78A    -----
HSP70         -----
HSP70.4       -----

```

**Figure 9.5. Sequence alignments of the *T. brucei* HSP70 family members.**

**(A)**

LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	KVPVDPAMYNTTTKDAYKPYDPDAYRREQPTDDGNGYEKVPVDPAMYNTTTKDAYKPYDP MRHVDESHFLTTTHEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTTHEAYKPIDP PRHVDPDHFRSTTQDAYRPVDP SAYKRALPLEEEEDVGPRHVDPDHFRSTTQDAYRPVDP ** : :***:*** * *. **: : ** : :***:*** * **
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	DAYRREQPTDDGEGYEKVPVDPAMYNTTTKDAYKPYDPDAYRREQPTDDGEGYERVPVDP SEYRQKRTVGEEVTTDMRHVDESHFLTTTHEAYKPIDPSEYRQKRTVGEEVTTDMRHVDE SAYKRALPLEEEEDVGPRHVDPDHFRSTTQDAYRPVDP SAYKRALPLEEEEDVGPRHVD . **: : ** : :***:*** * *. **: : **
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	AMYNNTTKDAYKPYDPDAYRREQPTDDGEGYEKVPVDPAMYNTTTKDAYKPYDPDAYRRE SHFLTTTHEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTTHEAYKPIDPSEYRQK DHFRSTTQDAYRPVDP SAYKRALPLEEQEDVGPRHVDPDHFRSTTQDAYRPVDP : :***:*** * *. **: : ** : :***:*** * *. **: :
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	QPTDDGEGYERVPVDPAMYNTTTKDAYKPYDPDAYRREQPTDDGEGYERVPVDPAMYNTT RTVGEEVATDMRHVDESHFLTTTHEAYKPIDPSEYRQKRTVGEEVATDMRHVDESHFLTT LPQEEEDVGPRHVDPDHFRSTTQDAYRPVDP SAYKRALPLEEEEDVGPRHVDPDHFRST : ** : :***:*** * *. **: : ** : :*
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	TKDAYKPYDPDAYRREQPTDDGEGYERVPVDPAMYNTTTKDAYKPYDPDAYRREQPTDDG THEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTTHEAYKPIDPSEYRQKRTVGEE TQDAYRPVDP SAYKRALPLEEEEDVGPRHVDPDHFRSTTQDAYRPVDP SAYKRALPQEEE *:***:*** * *. **: : ** : :***:*** * *. **: :
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	EGYERVPVDPAMYNTTTKDAYKPYDPDAYRREQPTDDGNGYEKVPVDPAMYNTTTKDAYK VATDMRHVDESHFLTTTHEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTTHEAYK EDVGPRHVDPDHFRSTTQDAYRPVDP SAYKRALPLEEQEDVGPRHVDPDHFRSTTQDAYR ** : :***:*** * *. **: : ** : :***:***
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	PYDPDAYRREQPTDDGNGYEKVPVDPAMYNTTTKDAYKPYDPDAYRREQPTDDGEGYEK PIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTTHEAYKPIDPSEYRQKRTIGEEVATDMR PVDP SAYKRALPLEEEEDVGPRHVDPDHFRSTTQDAYRPVDP SAYKRALPLEEEEDVGPR * ** . **: : ** : :***:*** * *. **: : *
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	PVDPAMYNTTTKDAYKPYDPDAYRREQPTDDGEGYERVPVDPAMYNTTTKDAYKPYDPDA HVDESHFLTTTHEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTTHEAYKPIDPSE HVDPDHFRSTTQDAYRPVDP SAYKRALPLEEEEDVGPRHVDPDHFRSTTQDAYRPVDP SAYKRALPLEEEEDVGPRHVDPDHFRSTTQDAYRPVDP ** : :***:*** * *. **: : ** : :***:*** * **.
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	YRREQPTDDGEGYEKVPVDPAMYNTTTKDAYKPYDPDAYRREQPTDDGEGYERVPVDPAM YRQKRTIGEEVTTDMRHVDESHFLTTTHEAYKPIDPSEYRQKRTIGEEVATDMRHVDESH YKRALPLEEEEDVGPRHVDPDHFRSTTQDAYRPVDP SAYKRALPLEEEEDVGPRHVDPDH *: : : ** : :***:*** * *. **: : **
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	YNTTTKDAYKPYDPDAYRREQPTDDGEGYEKVPVDPAMYNTTTKDAYKPYDPDAYRREQ FLTTTHEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTTHEAYKPIDPSEYRQKRT FRSTTQDAYRPVDP SAYKRALPLEEEEDVGPRHVDPDHFRSTTQDAYRPVDP SAYKRALPLEEEEDVGPRHVDPDHFRSTTQDAYRPVDP : :***:*** * *. **: : ** : :***:*** * *. **: :
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	TDDGEGYERVPVDPAMYNTTTKDAYKPYDPDAYRREQPTDDGEGYEKVPVDPAMYNTTTK VGEEVATDMRHVDESHFLTTTHEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTTH QEEEDVGPRHVDPDHFRSTTQDAYRPVDP SAYKRALPLEEEEDVGPRHVDPDHFRSTTQ : ** : :***:*** * *. **: : ** : :***:
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	DAYKPYDPDAYRREQPTDDGEGYEKVPVDPAMYNTTTKDAYKPYDPDAYRREQPTDDGEG EAYKPIDPSEYRQKRTVGEEATGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEV DAYRPVDP SAYKRALPLEEEEDVGPRHVDPDHFRSTTQDAYRPVDP SAYKRALPLEEQED :***:*** * *. **: : * *** :.:*** ::* * * * *: : *
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	YEKVPVDPAMYNTTTKDAYKPYDPDAYRRVDEVPSELPLVLALESDRAFSTPAETHVGG TGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREE---VVKVSRRGG-----D VGPRHVDPDHFRSTTQDAYRPVDP SAYKRESVVKD---VRAVNVRAH-----Y *** :.:*** ::* * * *: * * : . *
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	EGTAEATPSHVDVEQKPNPRMRAQADPAFYESIYAKDYKPAVPTAEAVPAPADVSEPA DGHFRTT-----A-KDSYCRYDSSCYKSSCRESSF---GDESVM PDTLRV-----S-HESYKSVDSAYKRESPV--V---KDLRAVNV . :.: : * :. ** *
LmjF.05.0380 Tb927.10.10360 TcCLB.511633.79	EKPDPRRRRAQADPAFYESIYAKDYKPAVPTAEAVPAPADVSEPAEKPDPMRRAQADL RH-----SFTASSEADA-----IV----- RH-----AYPDTLRVSHESYKLLN-----VASTRDGL----- .: : .:

LmjF.05.0380	SSYQSTYAKDYGCSGDGAPASVVGKPVVPLRIPSEGKSNKKNGTARTSVRAPIAGRQKER
Tb927.10.10360	-----VKKVVKGKSDGGVT-SKKVDVFTDMKCVASGESNGGAAAV---AAT---T-----
TcCLB.511633.79	-----SRAVCHRISDGKAA-QYGESSFSS--FVSNGDRNGTDGASSSSCRGSA---R-----
	*** : . :.*. * .:: :
LmjF.05.0380	SSVPAAKAHAPVSPPASAEAEPTKKTLPVAPAKTGEPAEPRPRPKGSYETEYRKNYRAYS
Tb927.10.10360	ECAAAA-----VGD-----Q---PTGNGVRVSK
TcCLB.511633.79	ACFGKS-----SSE-----VFESNFQTPKGT
	. : .: . :
LmjF.05.0380	NKHTPVETPRRTQACVHHVDPAFYTTTCSA-----
Tb927.10.10360	----WKKAPNVGYSCPHVDADMYVSTAHRDFKAHGASKPYMPKAAPTVKQSSISMQGV
TcCLB.511633.79	----DGHFSSKGYFCPCHTDPEMYRSTSHADYKAHHKDAYSRPYLKLPLDRKFPLERRDFL
	. * *. * :* :*.
LmjF.05.0380	-----AYVAPAAAPARSFAFTRPTVSVRHVDPSMYVSTSRAAYA-----
Tb927.10.10360	SEYRNRFRPEVQCPRPIGKPAVQVRHVDPSMYVTNTRATFVNHSGRKV
TcCLB.511633.79	SEYRKNFLRPEPQSLSRPVAASTVTVRHVDPSVYTTTNQAVFKDHWWKF-
	: : * . :* *****:*. :*. :*.

**(B)**

	:
Tb927.10.10360	RTVGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVATDMRHVDESHFLT
Tb927.10.10280	---G-----TKGLEADYKWNDPQYPQORDKERAENVDCIDIKNKKKSLPH
	* . . ** **.* *: . ** : :.: *
Tb927.10.10360	THEAYKPIDPSEYRQKRTVGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEE
Tb927.10.10280	NT-----ENLYMKQ-----QKANKQNYKEILSTMKESYAPIDPAAYKRDMLREEE
	. * : : . . : .*: * :*: * *: :. **
Tb927.10.10360	VATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVATDMRHVDESHFLTTHTEAYK
Tb927.10.10280	EVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDMLREEEEVTGPRHVDPSHFRSTTAESYA
	.*. **** * :** *.* ****: *: . ** .*. **** * :* *.*
Tb927.10.10360	PIDPSEYRQKRTVGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTIGEEVATDMR
Tb927.10.10280	PIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPR
	****: *: . ** .*. **** * :** *.* ****: *: . ** .*. *
Tb927.10.10360	HVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVATDMRHVDESHFLTTHTEAYKPIDPSE
Tb927.10.10280	HVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAA
	*** * :** *.* ****: *: . ** .*. **** * :** *.* ****:
Tb927.10.10360	YRQKRTIGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTIGEEVATDMRHVDESH
Tb927.10.10280	YKRDVLRREEEEATGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEATGPRHVDPSH
	*: . ** **.* **** * :** *.* ****: *: . ** **.* **** *
Tb927.10.10360	FLTTHTEAYKPIDPSEYRQKRTVGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRT
Tb927.10.10280	FRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDV
	* :.* *.* ****: *: . ** .*. **** * :** *.* ****: *: .
Tb927.10.10360	IGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVATDMRHVDESHFLTTH
Tb927.10.10280	REEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTA
	** .*. **** * :** *.* ****: *: . ** .*. **** * :**
Tb927.10.10360	EAYKPIDPSEYRQKRTVGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVA
Tb927.10.10280	ESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEV
	*:* ****: *: . ** .*. **** * :** *.* ****: *: . ** .
Tb927.10.10360	TDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTHTEAYKPI
Tb927.10.10280	TGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPS-----
	*. **** * :** *.* ****: *: . ** .*. **** *
Tb927.10.10360	DPSEYRQKRTVGEEVTTDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVATDMRHV
Tb927.10.10280	-----HFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPS
	** :** *.* ****: *: . ** .*. ***
Tb927.10.10360	DESHFLTTHTEAYKPIDPSEYRQKRTVGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYR
Tb927.10.10280	DPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYK
	* ** :** *.* ****: *: . ** .*. **** * :** *.* ****: *

Tb927.10.10360	QKRTVGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFL
Tb927.10.10280	RDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFR :. ** .*. **** ** :** *: * ****: *::. ** .*. **** **
Tb927.10.10360	TTTHEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVG
Tb927.10.10280	STTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLR :** *: * ****: *::. ** .*. **** ** :** *: * ****: *::.
Tb927.10.10360	EEVTTDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTHTEA
Tb927.10.10280	EEEVTTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTAES ** .*. **** ** :** *: * ****: *::. ** .*. **** ** :** *:
Tb927.10.10360	YKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTIGEEVATD
Tb927.10.10280	YAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTG * ****: *::. ** .*. **** ** :** *: * ****: *::. ** .*.
Tb927.10.10360	MRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTHTEAYKPIDP
Tb927.10.10280	PRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDP **** ** :** *: * ****: *::. ** .*. **** ** :** *: * ****
Tb927.10.10360	SEYRQKRTIGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVATDMRHVDE
Tb927.10.10280	AAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVD : *::. ** .*. **** ** :** *: * ****: *::. ** .*. ****
Tb927.10.10360	SHFLTTHTEAYKPIDPSEYRQKRTIGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQK
Tb927.10.10280	SHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRD *** :** *: * ****: *::. ** .*. **** ** :** *: * ****: *::.
Tb927.10.10360	RTVGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVATDMRHVDESHFLT
Tb927.10.10280	VLREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRST ** .*. **** ** :** *: * ****: *::. ** .*. **** ** :*
Tb927.10.10360	THEAYKPIDPSEYRQKRTIGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEE
Tb927.10.10280	TAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLR--- * *: * ****: *::. ** .*. **** ** :** *: * ****: *::.
Tb927.10.10360	VATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTIGEEVATDMRHVDESHFLTTHTEAYK
Tb927.10.10280	-----EEEVTTGPRHVDPSHFRSTTAESYA ** .*. **** ** :** *: *
Tb927.10.10360	PIDPSEYRQKRTVGEEVATDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVATDMR
Tb927.10.10280	PIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPR ****: *::. ** .*. **** ** :** *: * ****: *::. ** .*. *
Tb927.10.10360	HVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTHTEAYKPIDPSE
Tb927.10.10280	HVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAA *** ** :** *: * ****: *::. ** .*. **** ** :** *: * ****:
Tb927.10.10360	YRQKRTVGEEVTTDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVATDMRHVDESH
Tb927.10.10280	YKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSH *::. ** .*. **** ** :** *: * ****: *::. ** .*. **** **
Tb927.10.10360	FLTTHTEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTHTEAYKPIDPSEYRQKRT
Tb927.10.10280	FRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLR * :** *: * ****: *::. ** .*. **** ** :** *: * ****: *::.
Tb927.10.10360	VGEEVTTDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTH
Tb927.10.10280	REEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTA ** .*. **** ** :** *: * ****: *::. ** .*. **** ** :**
Tb927.10.10360	EAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVA
Tb927.10.10280	ESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEV *: * ****: *::. ** .*. **** ** :** *: * ****: *::. ** .
Tb927.10.10360	TDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVATDMRHVDESHFLTTHTEAYKPI
Tb927.10.10280	TGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPI *. **** ** :** *: * ****: *::. ** .*. **** ** :** *: * **
Tb927.10.10360	DPSEYRQKRTVGEEVTTDMRHVDESHFLTTHTEAYKPIDPSEYRQKRTVGEEVATDMRHV
Tb927.10.10280	DPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEATGPRHV ** : *::. ** .*. **** ** :** *: * ****: *::. ** ** . **
Tb927.10.10360	DESHFLTTHTEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTHTEAYKPIDPSEYR
Tb927.10.10280	DPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYK * ** :** *: * ****: *::. ** .*. **** ** :** *: * ****: *

Tb927.10.10360	QKRTVGEEVTTDMRHVDESHFLTTHAEYKPIDPSEYRQKRTIGEEVATDMRHVDESHFL
Tb927.10.10280	RDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFR
	:. ** .*. **** ** : ** : * **** : * : . ** .*. **** **
Tb927.10.10360	TTTHEAYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTHAEYKPIDPSEYRQKRTIG
Tb927.10.10280	STTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLR
	: ** : * **** : * : . ** .*. **** ** : ** : * **** : * : .
Tb927.10.10360	EEVTTDMRHVDESHFLTTHAEYKPIDPSEYRQKRTIGEEVATDMRHVDESHFLTTHAE
Tb927.10.10280	EEVTTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAES
	** .*. **** ** : ** : * **** : * : . ** .*. **** ** : ** :
Tb927.10.10360	YKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTHAEYKPIDPSEYRQKRTVGEEVATD
Tb927.10.10280	HAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTG
	: **** : * : . ** .*. **** ** : ** : * **** : * : . ** .*.
Tb927.10.10360	MRHVDESHFLTTHAEYKPIDPSEYRQKRTVGEEVTTDMRHVDESHFLTTHAEYKPIDP
Tb927.10.10280	PRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDPSHFRSTTAESYAPIDP
	**** ** : ** : * **** : * : . ** .*. **** ** : ** : * ****
Tb927.10.10360	SEYRQKRTVGEEATGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEVTGPRHVDP
Tb927.10.10280	AAYKRDVLRREEEEATGPRHVDPSHFRSTTAESYAPIDPAAYKRDVLRREEEEATGPRHVDP
	: * : . *****
Tb927.10.10360	SHFRSTTAESYAPIDPAAYKRDVLRREEVVKVSRGGDDGHFRTTAKDSYCRYDSSCYKSS
Tb927.10.10280	SHFRSTTAESYAPIDPAAYKRDVLRREEVVKVSSRGGDDGHFRTTAKDSYCRYDSSCYKSS
	*****
Tb927.10.10360	CRESSFGDESVMRHSFTASSEADAIVVKVVKGKSDGGVTSKKVDVFTDMKCVASGES
Tb927.10.10280	CRESSFGDESVISKRHSFTASSVADAIVVKVVKGKSDGGVTSKKVDVFTDMKCVASGES
	***** : * *****
Tb927.10.10360	NGGAAAATAATTECAAAVGDQPTGNGVRVSKWKKAPNVGYSCPCHVADADMYVSTAHRDF
Tb927.10.10280	NGGAAAATAATTECAAAVGDQPTGNGVRVSKWKKAPNVGYSCPCHVADADMYVSTAHRDF
	*****
Tb927.10.10360	KAHGASKPYMPKAAPTVKQSSISMQGSSEYRNRFIRPEVQCPRPIGKPAVQVRHVDPS
Tb927.10.10280	KAHGASKPYMPKAAPTVKQSSISMQGSSEYRNRFIRPEVQCPRPIGKPAVQVRHVDPS
	*****
Tb927.10.10360	MYVTNTRATFVNHSGRKV
Tb927.10.10280	MYVTNTRATFVNHSGRKV
	*****

**Figure 9.6 (A) Sequence alignments of the microtubule-associated proteins in trypanosomatids. (B) Sequence alignment shows high similarity of two microtubule-associated proteins in *T. brucei*.**

### (A)

TbBBP590	MPFAPQGPVMDLYNSDEGDLAIRSEMWMCADSEKMFVKVAGRGSKNIVSIENGHFFVYS
T.cruzi_TCDM_02199-t26	-----
TbBBP590	TCGKQSRLIKDVFFHSVKRTASFVHDPKVSGASTVSGMQTAASSSSSSARSRRRADVQV
T.cruzi_TCDM_02199-t26	-----

TbBBP590 T.cruzi_TCDM_02199-t26	PPYYLVLEFVRDKTVAESTNLAKREHVVLCTDNVQDFQYWQQFVGRYKCQLNRTQRPDE -----
TbBBP590 T.cruzi_TCDM_02199-t26	SVVGGRRSGKGCVKASASHKECDDSGSDETNLDISTLTMKLDELQEWKRIAVDLADAVASS -----
TbBBP590 T.cruzi_TCDM_02199-t26	GAFFNCRRESLSNTEPNVEFSLKYVEKLLGELRKGVDCLNLTNETQPPLPLVKDLKQQQL -----
TbBBP590 T.cruzi_TCDM_02199-t26	ELWNSLYDRTFSDDRTSEARGLFTKERQQVVRVQRMENI IARERTEGCTKLSPVVSRREL -----
TbBBP590 T.cruzi_TCDM_02199-t26	TSPEEWLQLPCLSTLSDADSAQNVLGGQREGGVLDSAITLLSSSTVHADAPQEVWQRE -----
TbBBP590 T.cruzi_TCDM_02199-t26	NLSSANEGDMFAELRAAHRNLGLRSAPQHEFHQLEREKKNTERLAREGLRAYTEATIVLL -----
TbBBP590 T.cruzi_TCDM_02199-t26	QACYSEYVTVVECVLSAISGGVDTRGERRLVTAYAHLQTVAAALDEARRAATAKQTILENL -----
TbBBP590 T.cruzi_TCDM_02199-t26	VNELKDLQKTRDEEINALRKTFFRAKDGWESDLAVLQTKLRTLQNTMTQTTVVPSDTVS -----
TbBBP590 T.cruzi_TCDM_02199-t26	FPPLERALCGGESSQELGRDSGEDKVSNTKIESVDKASYDECDEADSDSIRRWATRLTC -----
TbBBP590 T.cruzi_TCDM_02199-t26	SHGEQRGHI IAFRLERFFHWTLENVVPLCSGADTNCLLELVMRALSNHITLLQAAGEVFG -----
TbBBP590 T.cruzi_TCDM_02199-t26	TQDVKVSVDNNGVRESRDLCSIRDLTQHSTLKLHLVGKYIALTENEEDSPVLRCDLEK -----
TbBBP590 T.cruzi_TCDM_02199-t26	QLQLHQFNTATLDHINSLGTDNMTIKRIQDMIQNEEQLLRIQEITGSKREICTVQHV -----
TbBBP590 T.cruzi_TCDM_02199-t26	QKHEELRQMSELINRLDQSRMHQEKDLTTQNYEFTQMNTSVKAAEII ENTTKKYNELQ -----
TbBBP590 T.cruzi_TCDM_02199-t26	LRIDFDKIEFNERLQIVHEYNEHPQHYHNNKINKQIEELNEIEKQQLNIGHQDKKEKK -----
TbBBP590 T.cruzi_TCDM_02199-t26	TERNSSTRAQCLOGIESALAVLGTTEAEMPLASAIQRAAEEAKVREDTLKQELDEAMGRE -----
TbBBP590 T.cruzi_TCDM_02199-t26	TALEEQLRNLSQATIALQGTLYGGEMPTGNDGSALDVGHAAMSALERMREDAAKASGVM -----
TbBBP590 T.cruzi_TCDM_02199-t26	AAMQEGEMSVPLLAKCHHMAETTNALRDSEKHYRALVGEDVKKDSGSTGTDDSSTRAQCL -----
TbBBP590 T.cruzi_TCDM_02199-t26	QGIESALAVLGTTEAEMPLASAIQRAAEEAKVREDTLKQELDEAMGRETALEEQLRNLSQ -----
TbBBP590 T.cruzi_TCDM_02199-t26	ATIALQGTLYGGEMPTGNDGSALDVGHAAMSALERMREDAAKASGVMTAAMQEGEMSVPL -----
TbBBP590 T.cruzi_TCDM_02199-t26	LAKCHHMAETTNALRDSEKHYRALVGEDVKKDSGSTGTDDSSTRAQCLOGIESALAVLGT -----
TbBBP590 T.cruzi_TCDM_02199-t26	TEAEMPLASAIQRAAEEAKVREDTLKQELDEAMGRETALEEQLRNLSQATIALQGTLYGG -----
TbBBP590 T.cruzi_TCDM_02199-t26	EMPTGNDGSALDVGHAAMSALERMREDAAKASGVMTAAMQEGEMSVPLLAKCHHMAETTN -----

TbBBP590 T.cruzi_TCDM_02199-t26	ALRDSEKHYRALVGEDVKKDSGSTGTDDSSSTRAQCLQGIESALAVLGTTEAEMPLASAIQ -----
TbBBP590 T.cruzi_TCDM_02199-t26	RAAEEAKVREDTLKQELDEAMGRETALEEQLRNLSQATIALQGTLYGGEMPTGNDGSALD -----
TbBBP590 T.cruzi_TCDM_02199-t26	VGHAAMSALERMREDAAKASGVMTAAMQEGEMSVPLLAKCHHMAETTNALRDSEKHYRAL -----
TbBBP590 T.cruzi_TCDM_02199-t26	VGEDVKKDSGSTGTDDSSSTRAQCLQGIESALAVLGTTEAEMPLASAIQRAAEEAKVREDT -----
TbBBP590 T.cruzi_TCDM_02199-t26	LKQELDEAMGRETALEEQLRNLSQATIALQGTVCEDGADVSCDISTFDLCSVAFSLGVI -----
TbBBP590 T.cruzi_TCDM_02199-t26	ESAFECISISREVSLLKKQVEEAHKREVALRSRLKSLASAATAVQSTLDGNASPERVSPKES -----
TbBBP590 T.cruzi_TCDM_02199-t26	SICSGDEVGDGDAFVDPFVAVQRAAEEKACRNAMEEKLKEGAARIDAVEGRLLQLSRSI -----
TbBBP590 T.cruzi_TCDM_02199-t26	ICLHSLKFDAEIGNGEERAVELSEALVAADAIENAFECISISREVSLLKKQVEEAHKREV -----
TbBBP590 T.cruzi_TCDM_02199-t26	ALKTRLKCLVGAATAVQCTLEGGNYSCTADDPLSLSGCDMSASTNLDDSVLGSVAAIQRA -----
TbBBP590 T.cruzi_TCDM_02199-t26	AEEAKVREDTLKQELDEAMGRETALEEQLRNLSQATIALQGTLYGGEMPTGNDGSALDVG -----
TbBBP590 T.cruzi_TCDM_02199-t26	HAAMSALKRMREDAAKASGVMTAAMQEGEMSVPLLAKCHHMAETTNALRDSEKHYRALVG -----
TbBBP590 T.cruzi_TCDM_02199-t26	EDVKKDSGSTGTDDSSSTRAQCLQGIESALAVLGTTEAEMPLASAIQRAAEEAKVREDTLK -----
TbBBP590 T.cruzi_TCDM_02199-t26	QELDEAMGRETALEEQLRNLSQATIALQGTLYGGEMPTGNDGSALDVGHAAMSALERMRE -----
TbBBP590 T.cruzi_TCDM_02199-t26	DAAKASGVMTAAMQEGEMSVPLLAKCHHMAETTNALRDSEKHYRALVGEDVKKDSGSTGT -----
TbBBP590 T.cruzi_TCDM_02199-t26	DDSSSTRAQCLQGIESALAVLGTTEAEMPLASAIQRAAEEAKVREDTLKQELDEAMGRETA -----
TbBBP590 T.cruzi_TCDM_02199-t26	LEEQLRNLSQATIALQGTLYGGEMPTGNDGSALDVGHAAMSALKRMREDAAKASGVMTAA -----
TbBBP590 T.cruzi_TCDM_02199-t26	MQEGEMSVPLLAKCHHMAETTNALRDSEKHYRALVGEDVKKDSGSTGTDDSSSTRAQCLQG -----
TbBBP590 T.cruzi_TCDM_02199-t26	IESALAVLGTTEAEMPLASAIQRAAEEAKVREDTLKQELDEAMGRETALEEQLRNLSQAT -----
TbBBP590 T.cruzi_TCDM_02199-t26	IALQGTLYGGEMPTGNDGSALDVGHAAMSALERMREDAAKASGVMTAAMQEGEMSVPLLA -----
TbBBP590 T.cruzi_TCDM_02199-t26	KCHHMAETTNALRDSEKHYRALVGEDVKKDSGSTGTDDSSSTRAQCLQGIESALAVLGTTE -----
TbBBP590 T.cruzi_TCDM_02199-t26	AEMPLASAIQRAAEEAKVREDTLKQELDEAMGRETALEEQLRNLSQATIALQGTLYGGEM -----
TbBBP590 T.cruzi_TCDM_02199-t26	PTGNDGSALDVGHAAMSALERMREDAAKASGVMTAAMQEGEMSVPLLAKCHHMAETTNAL -----



[illegible]

[illegible]

TbBBP590	LKTSQQANEEEFARVQEQMQELRHAVRKKVEDDRKVEQSLRDLNDHLDMQARELAMKYRA
T.cruzi_TCDM_02199-t26	LIHERRQTEEEQARVEDHVQELRRVVQRKIEEDRRVEQSLKELDRHLDMQARELAMKYRS
	* .:: .*** ***::::***:.*:*.***:***:.*.*****:
TbBBP590	DQEAISRQFAELRGSIRRVVIPTQEQLGTPSPKYDFVYPRE-
T.cruzi_TCDM_02199-t26	DHDIIVRQFTELRGTIIRRTIKPPRSHSTAAT-TFGSASHRTF
	*.: * ***:***:***.: * .:: : : .:. . *

## (B)

TbBBP590	MPFAPQGPMVDLYNSDEGDLLAIRSEMWMCADSEKMFKVAGRGSKNIVSIENGHFFVYS
LmjF.36.5250	-----MSFMEEMSDAPVSITGEMWMCADSAGPLKVGAGRGLKNLITIEKDRFCAYS
	*.: .: . . :.* .***** :***** **::*:.* .**
TbBBP590	TCGKQSRLIKDVPFHSVKTASFVHDPKVSGASTVSGMQTAASSSSSSARSRRRADVQV
LmjF.36.5250	SRSEQCKQVKQIAFRSMKCVAWFAHHKPPGREPAVST-----TAATLRRGARKDPLA
	: .*:.* :*:.* *:.* * .*. * . . . :::: *. * * .
TbBBP590	PPYYLVLEFVRDKTVAESTNLAKREHVVLCTDNVQDFQYWQQFVGRYKCQLNRTQRPDE
LmjF.36.5250	ATYYLVLEFLRDNTIPGTVSLGKRERIVLCTDDQQDFS IWRKFAELYECPPVHEGFEVG
	*****:***: .:.*.***:*****: ***. *::*. *: * : :
TbBBP590	---SVVGRRSGKGCVKKSASHKECDDSGSDETNLDISTLTMKLDDELQEWKRIAVDLADAV
LmjF.36.5250	RDKKALRAQAPAGGARDHDDTTSSDDEGYEGRGLDGG-ASVSRPLDLWKNRCVALLNDF
	.:. :. . *. . . .*. * : .** .: :. *: ** . * * : .
TbBBP590	ASSG-----AF---PNCRESLSNTEPNVEFSLKYVEKLLGE-----LRKGV
LmjF.36.5250	AHLSNPSLQVQSADVFDTAGCGGLDW---DAHVESVLHAIQDMRPSMKVVPAPSTRGA
	* . . * . * . : : ** * : * : : : *
TbBBP590	DCRLNTLNETQ-----PP-----LPLVKDLKQLELWNSLYDR--TFS
LmjF.36.5250	AAGPSSLLSSPSIAPSSIHPCAATKGTMETDFDEMQLASNLQVELEQFQVVAEEVGALA
	. .: * .: * *..*: : ** : : . : :
TbBBP590	DDRTSEARGLFTKERQQVRVQRMENIIARETEGCTKLSPVVSRRELVT-----SPE
LmjF.36.5250	EGRELKTTTTLPD-ALHAYVERLRKMPTPA-----DPADSVAAQREDKTAANASAAAADL
	:.* : : : . . *:.*: : : . *. : ** * :
TbBBP590	EWLQLPCLSTLSDADSAQNLVGGQREGGVLDGSAITLLSSSTVHADAPQEVWQERENLSS
LmjF.36.5250	AKLESACVNAVRSFT-----DKVHTSALPLSQEWNREKAEQQ
	*: *::: . . . * .: *::: ..
TbBBP590	A-----NEGDMFAELRA--
LmjF.36.5250	ERAALPQRVETIVAVLIDELTRAAADVSALKTSLATSMAPTASPATNVPLELSELRRVL
	* : : **
TbBBP590	AHRNLGLRSAPQHEFHQLEREKKNTERLAREGLRAYTEATIVLLQACYSEYVTVVECVLS
LmjF.36.5250	NQKELEVETLQQRHFHASRHTKQ-LSRRLEELVFSYSDVAHDMRLSYREYASLRNLFLR
	:*: .: *.** .: *: . * . * : :*: : : : *: * ** : : . *
TbBBP590	AISGGVD-----TRGERRLVTAYAHLQTVAALD
LmjF.36.5250	ALRGEVPLDNSTAAEASLGESGSGGTSEARCSGHLPGSFTGVQKLVISAHLEAIQALE
	*: * * * : ** : **::: *
TbBBP590	EARRAATAKQTLLENLVNELKDLQKTRDEEINALRKTFFRKDGWESDLAVLQTKLRTLQ
LmjF.36.5250	QERIAARERQRELEKVMKEVLRQERDEEIANLKCNFLAAGEAWEKDVAVLQAKLTALQ
	: * ** : * **:::***: *: ***** *: .*: * : .*. *:***: * * **
TbBBP590	NTMTQTTVVVPSDTVSPFLERALCGGESSQELGRDSGEDKV-SNTKIESVDKASYDECD
LmjF.36.5250	SAIPRGALLLPASSLAAHTAVKA---GI-AQKLQSTQPHTPTAPADTRESTETV---TPM
	.: : : :*: : : : * * :*: * . . . . **::..
TbBBP590	EADSDSIRRWATRLTC--SHGEQRGHIIAFRLERFFHWTLENVVPLCSGADTN-CLEEL
LmjF.36.5250	LAESHKTLHKLTTLLATTPKFTGVEAAVVTQRLQAFYDWLEAVVPLSSTSVDGSLVEI
	*:.*::: : * * . . .: ** *:.*.* ***. * : : :.*:
TbBBP590	VMRALSNHITLLQAAGEVFGTQDVKVSDDNGVRESRDLCCSIRDHLHTQHSTLKHVLGKY
LmjF.36.5250	ITHVVHAYHALFQQISTITYATRS--GTDDTHV---DVLKLV---KNEHEKFEKLWGI-
	: : : : :*: . : : .*. * *: . : :.*.:*: *
TbBBP590	IALTENEEDSPVLRCDLEKQLQLHQFNATATLDHINSLGTDNMTIKRRIQDMIQNEEQQL
LmjF.36.5250	-----LEKHLIHGE-----DSS
	***: * : *

TbBBP590 LmJF.36.5250 LRIQEITGSKREICTVQHVQRKHEELRQMSELINRLDQSRMHQEKDLTTQNYEAFQMNT  
FHISGHQSSSERTALLFSSPTMEELDQLPSFLETLW-----  
::\*. \*. . . . : . \*\*\* \*: .::: \*

TbBBP590 LmJF.36.5250 SVKKAEEIIENTTKKYNELQLRIDFDKIEFNERLQIVHEY--NEHPQHYHNNKINKQI-  
-----TNATLYSQLKLRSRG--VPIATYLDLEKKSAQLEQQQLQHTQREKRSLS  
. . . . \*:\*:\*\* . : : \*\* :. : \*: \* \*: : . . .

TbBBP590 LmJF.36.5250 ---EELNEIIEKQONLIGHQDKKEKKTURNSSTRACQLQGIESALAVLGTTEAE-MPLAS  
ALMATFQDDKHHTSSLIGHAN-----SALAALGVGAPASESAAE  
: : . : ..\*\*\*\* : \*\*\*\*.\*. \*

TbBBP590 LmJF.36.5250 AIQRAAEEAKVREDTLKQELDEAMGRET-----ALEEQLRNL-----SQATI  
AIARAAADAAERERALARAEVAAMEERLEAANAQQQLQADRDEQLRDADERLAETAAAHQ  
\*\* \*\* \*: \* \*: \*: \* :\*\*\*\*: : \*

TbBBP590 LmJF.36.5250 ALQGTLYGGEMPTGNDGSALDVGHAA---MSALERMREDAKASGVT---AAMQEGEM  
ELSGSVADAL-----AALGVGAPASESAAEAIARAAADAAERERALARAEVAAMEER-L  
\*.\*: . . :\*.\*\* \* .\*: \* \*\*\*: . . : \*\*\*:\* :

TbBBP590 LmJF.36.5250 SVPLLAKCHHMAETTNALRDSEKHYRALVGEDVKKDSGSTGTDDSSTRAQCLQGIESALA  
EAANAQQQLQADRDEQLRDADERLAET-----AAAHQELSGSVADALA  
. . : : \*: : \*\*\*\*: : : : : : . : .\*\*\*

TbBBP590 LmJF.36.5250 VLGTEAE-EMPLASAIQRAAEEAKVREDTLKQELDEAMGRET-----ALEEQLR  
ALGVGAPASESAAEAIARAAADAAERERALARAEVAAMEERLEAANAQQQLQADRDEQLR  
.\*. . \*.\*\* \*\* \*: \* \*: \*: \* :\*\*\*\*

TbBBP590 LmJF.36.5250 NL-----SQATIALQGTLYGGEMPTGNDGSALDVGHAA---MSALERMREDAKASG  
DADERLAETAAAHQELSGSVADAL-----AALGVGAPASESAAEAIARAAADAAERER  
: : \* \*.\*: . . :\*.\*\* \* .\*: \* \*\*\*: .

TbBBP590 LmJF.36.5250 VMT---AAMQEGEMSVPLLAKCHHMAETTNALRDSEKHYRALVGEDVKKDSGSTGTDDSS  
ALRAEVAAMEER-LEAANAQQQLQADRDEQLRDADERLAET-----AA  
. : \*\*\*:\* :. . : : \*: : \*\*\*\*: : : :

TbBBP590 LmJF.36.5250 TRAQCLQGIESALAVLGTTEAE-EMPLASAIQRAAEEAKVREDTLKQELDEAMGRETALEE  
AHQELSGSVADALAALGVGAPASESAAEAIARAAADAAERERALARAEVAAM-----EE  
: : : . : .\*\*\*.\*. . \*.\*\* \*\* \*: \* \*: \*: \*

TbBBP590 LmJF.36.5250 QLRNLSQATIALQGTLYGGEMPTGNDGSALDVGHAAASALERMREDAKASGVMTAAMQE  
RLEAANAQQQLQ-----ADRDEQLLD-----ADERL-----  
:\*. . \*\* :..\* . \*\* \* \*\*:

TbBBP590 LmJF.36.5250 GEMSVPLLAKCHHMAETTNALRDSEKHYRALVGEDVKKDSGSTGTDDSSTRAQCLQGIES  
-----AETAAA-----HQELSGSVAD  
\*\*\*: \* : : . : .

TbBBP590 LmJF.36.5250 ALAVLGTTEAE-EMPLASAIQRAAEEAKVREDTLKQELDEAMGR-----ETALEE  
ALAALGVGAPASESAAEAIARAAADAAERERALARAEVAAMEERLEAANAQQQLQADRDE  
\*\*\*.\*. . \*.\*\* \*\* \*: \* \*: \*: \* : : : \*

TbBBP590 LmJF.36.5250 QLRNLS-----QATIALQGTLYGGEMPTGNDGSALDVGHAA---MSALERMREDAK  
QLRNADERLAETAAAHQELSGSVADAL-----AALGVGAPASESAAEAIARAAADAAE  
\*\*\*. . \* \*.\*: . . :\*.\*\* \* .\*: \* \*\*\*:

TbBBP590 LmJF.36.5250 ASGVMTAAMQEGEMSVPLLAKCHHMAETTNALRDSEKHYRALVGEDVKKDSGSTGTDDSS  
RERA-----LRAEVAAMEERLEAANAQQQLQADRDEQLRNADE--RLAETA  
. . \* \*: \* \*: \* . . : : : \* .\*: : . : :

TbBBP590 LmJF.36.5250 TRAQCL-QGIESALAVLGTTEAE-EMPLASAIQRAAEEAKVREDTLKQELDEAMGR----  
AAHQELSGSVADALAALGVGAPASESAAEAIARAAADAAERERALARAEVAAMEERLEAAN  
: \* \* . : .\*\*\*.\*. . \*.\*\* \*\* \*: \* \*: \*: \*

TbBBP590 LmJF.36.5250 -----ETALEEQLRNLSQATIALQGTVCEDGADVSCDISTFDLCSVAFLSLGVIESAFE  
AQQQLQADRDEQLRDCVRERCAYVSQVCLRELEAKEAEARFDVVG-----AY  
: : :\*\*\*\*: : \* . \*\* :. . : \*\*:

TbBBP590 LmJF.36.5250 CSISREVSLKK-----QVEEAHKREVALRSRLKSLASAATAVQSTLD-GN  
CALSPYVSIASSIIGKLLDIRGMVGGDVDEIVGRVRALVAELNGLRAFASEAHEKVVWYE  
\*: \* \*\*: . :\*: \* \*\* :\*: \* : : : :

TbBBP590 LmJF.36.5250 ASPERVSPKES----SICSGDEVGDGDAFVDPFVAVQRAAEESKACRNALEEKKEGAA  
DVRVRVPLKKLSDFVRKRDGTDECEQREEFSVSFLA-----ETLRSSFR  
\*\* \*: . . \*\* : : \* \*: \* \*.\*:.

TbBBP590 RIDAVEGRLLQLSRSIICLHSLKLFDAEIGNGEERAVELSELALVAADAIENAFECISISR

LmjF.36.5250 VLQCLDSEAMALQESIFSMTALLKE--CGTALNE-----SVTNLQKALPCAASK  
: : : : : \* . \* : : \* : \* : : : \* : \* :

TbBBP590 EVSLKKKVEEAHKREVALKTRLKCLVGAATAVQCTLEGGNYSTADDPLSLSGCDMSAST  
LmjF.36.5250 EDTSPSQSHST--LPVEWGLRVVC-----RRLGAT-----  
\* : . : . : \* \* : \* \* .

TbBBP590 NLDDSVLGSVAIIQRAAEEAKVREDTLKQELDEAMGRETALEEQLRNLSQATIAL-----  
LmjF.36.5250 -----VSEA----VHWQGELTNILKATYRILIATS  
: . \* . : : \* \* : \* \* :

TbBBP590 QGTL-YGGEMPTGNDGSALDVGHAAMSALKRMREDAAKASGVMTAAMQEGEMSVPLLAKC  
LmjF.36.5250 DGTRVHGAPVSDGDHPSF-----AEYSQLPRLAA-----QAC  
: \* : \* . : \* . \* \* \* \* :

TbBBP590 HHMAETNALRDSEKHYRALVGEDVKKDSGSTGTDDSSTRAQCLQGIESALAVLGTTEA-  
LmjF.36.5250 SELLSRR-FAAVEDELAA-VTQQ---RN-----DLTGTGTCCGSLCSAISVATPGRVG  
: : . : : \* . \* \* : . \* : . \* . : \* : \* .

TbBBP590 EMPLASAIQRAAEEA-----KVREDTLKQELDEAMGRETALEEQLRNLSQATIALQGT  
LmjF.36.5250 ALPATA-TSLTCDEAKSALRDTVLTSSILHDYAQSMEHTLAVLTPLASIAADTFAFSSS  
: \* : : . : : \* . \* . : : : \* : \* : \* : \* : \* : :

TbBBP590 LYGGEMPTGNDGSA-----LDVGHAAMSALERMREDAAKASGVMTAA-----  
LmjF.36.5250 FEM--LAAATEAVRARIKRMERSQLQHLHMSQCD--EEGNAKARGAMQDQLFALSEDGEA  
: : : : . : \* \* : . \* . \* . \*

TbBBP590 -----MQEGEMSVPLLAKCHHMAETNALRDSEKHYRALVGEDVKKDSGSTGTDDSSTR  
LmjF.36.5250 GELARNEQEEQAAPLAE-VTLQRQRWDAAEVLGRFAEDVGATVRAKNGSESQSVGNT-  
: : \* : \* : : : \* : \* : \* . \* . . \*

TbBBP590 AQCLQGIESALAVLGTTEAEMPLASAIQRAAEEAKVREDTLKQELDEAMGRETALEEQLR  
LmjF.36.5250 -EVLPLVRQAISLSASAK-----  
: \* : . \* : : : :

TbBBP590 NLSQATIALQGTLYGGEMPTGNDGSALDVGHAAMSALKRMREDAAKASGVMTAAMQEGEM  
LmjF.36.5250 ---AQEELEGRL-----SGILDACKR-G--  
\* \* : \* \* \* : \* : \*

TbBBP590 SVPLLAKCHHMAETTN---ALRDSEKHYRALVGEDVKKDSGSTGTDDSSSTR-AQCL---Q  
LmjF.36.5250 ---AAEHHWAEATAARMALDPEELHASVTTEGHS-----PEKMQRANACQMPIA  
\* \* \* \* : \* \* . \* . \* . : . : \* \*

TbBBP590 GIESALAVLGTTEAEMPLASAIQRA-AEEAKVRE-----  
LmjF.36.5250 TMEAAARAALQSTTQMQLVAARAHEEALDVGVRWDASESGVRAPAGEGERLGEEMYLVQ  
: \* : \* . \* : : \* \* : \* : \* . \* . \*

TbBBP590 ---DTLKQ-----ELDEAMGRETALEEQLRNLSQATIALQGTLYGGEM  
LmjF.36.5250 SRMDAMQDRPNEAERGSSQVMDAGVLLAAGQVRSLEDVMRVVASYLEYKTKS-----  
\* : : : \* \* \* : \* : \* : : :

TbBBP590 PTGNDGSALDVGHAAMSALERMREDAAKASGVMTAAMQEGEMSVPLLAKCHHMAETT---  
LmjF.36.5250 -----LE-----SKESRLRSEIVQDWA-----VMMVPLAVELRLLRRAIAGV  
\* : \* . \* : . : \* \* \* : : :

TbBBP590 -----NALRDSEKHYRALVGEDVKKDSGSTGTD-DSSTRAQCLQGIESALAVLGTTEAE  
LmjF.36.5250 DRPLPGTALADLPA---HGSSVRRELTEAGRSLGAAASAPCTEEAT-REALLGAMDQA  
\* \* \* \* . \* : : . \* : \* : \* : \* : :

TbBBP590 MPLASAIQRAAEEAKVREDTLKQELDEAM--GRETALEEQLRNLSQATIALQGT-----L  
LmjF.36.5250 CQLLL---GAPPER-----VLDTAAPHGTASSAEE-----CPAVTFAGAVSTTDV  
\* . \* : \* \* \* \* : \* \* : : \* : :

TbBBP590 YGGEMPTGND-----GSALDVGHAAMSALERMREDA--KASGVMTAA  
LmjF.36.5250 AGGTAPDKLADTVYAMIEERGDRAAANGTLCTGEAWYSAVARVTQTLCSALSSAGVAVSV  
\* \* \* . : \* . \* \* : \* : . : \* \* . :

TbBBP590 MQEGEMSVPLLAKCHHMAETNALRDSEKHYRALVGEDVKKDS---GSTGTDDSSTRAQC  
LmjF.36.5250 PADHGF-VVGHSSCVSPASQTSM----LDFLASAAEHVLYKTTDTTRLFEDWRESLQDC  
: : \* : . \* . \* . : \* . \* . \* : \* : :

TbBBP590 LQGIESALAVLGTTE--AEMPLASAIQRAAEEAKVREDTLKQELDEAMGRETALEEQLRN  
LmjF.36.5250 LSRTESSLPTLVAAVVEEASLRAASQLAAEEQQRHDEL---D-----ALRT  
\* . \* : \* . \* : : \* : \* \* \* : \* : \* \*

TbBBP590 LSQATIALQGTLYGGEMPTGNDGSALDVGHAAMSALERMREDAAKASGVMTAAMQEGEM-  
LmjF.36.5250 LLQ-----EASAVFSDTSPFLSMQ

```

* *                                     :*.*.:: :      .*

TbBBP590      SVPLLA KCHHMAETTNALRDSEKHYRALVGEDVKKDSGSTGTDDSSSTRAQCLQGIESALA
LmjF.36.5250  SPNFVEKCMAYAKECARLTALESELSSVLGAASAQRSS-----WEDASV
*   :: **   *:   *   *..   ::*   : * .   * . * .

TbBBP590      VLGTTEAEMPLASAIQRAAEEAKVREDTLRQELDEAMGRETALEEQLRNLSQATIALQGT
LmjF.36.5250  SLGDLAAEL-----QRDRVE-----LRE--VEVPKLGQL
**   **::   **   *   **::   .   ***

TbBBP590      VCEDGADVSC-----DISTFDLCSVAFSLSLGVIESAFECSSISREVS LKKQ
LmjF.36.5250  VCTLSVEVLLAEERAARTMLVAEEATSRHLHALCDPQLRELGLFVER--ALARAATAEAQ
**   ..:*   .   ** .   : .**:: .   ::* . : : *

TbBBP590      VEEAHKREVALR---SRLKSLASAATAVQS-----TLDG---NASPERVSPKESSIC
LmjF.36.5250  CAEKTAQHCSLLAAHESLVTMASGLVPVQEDCCRALVESEGDAACAAPTRFTSEALRIL
*   . : * .   *   : :*. .   **.   : *   * : * . : : *

TbBBP590      SGDEVGDGDAFVDPFVAV-----QRA-AEESKACRNALEEKLKEGAARIDAV--
LmjF.36.5250  E-----QEAVLVPRLVQAHGRLTATCAALRSDLDAKAACVQA---LADGEADLDIARS
.   : .:*   : * .   :   .   ** : *   * : * * : * .

TbBBP590      -----EGRLLQLSRSIICLHSLKLFDAEIGNGEER----AVELSELALVAADAIENAFEC
LmjF.36.5250  KSTELEAQLEHVLDIAEALGSLLPHEEASLRDASCLPRLVEAA-GQLAHTAAADAARCA
* . : * : : * . * * * * * . . :   ** :   * . : * : * . *

TbBBP590      SISREVS LKKKV EEAHKREVALKTRLKCLVGAATAVQCTLEGGNYSCTADDPLSLSGCDM
LmjF.36.5250  RLERDVCDGAAFA-----EQQRQCTAAESNAETLQETARRYECQEAEITQLTASAS
: . : * .   .   : : : * . . : . * .   .   . * . * : : . .

TbBBP590      SASTNLDDSVLGSVAAIQRAAEEAKVREDTLKQELDEAMGRETA-LEEQLRNLSQATIAL
LmjF.36.5250  KT---AAEAYAGIVTALCLSAESVP-----SAASAAAISEEQWAAV---VHAV
. :   : :   * * : * : **..   . * . : *   ***   :   . * :

TbBBP590      QGTLYGGEMPTGNDGSALDVGHAAAMSAL-ERMREDAAKASGVMTAAMQEGEMSVP LLA K-
LmjF.36.5250  QAAVEEGGRQTS-DAAELRVE---MRILQSRMAE---LKAAVLESEAREGNLLQRLNALM
* . : :   *   * . * . : * *   *   * . ** *   . : . : :   : : * : :   * *

TbBBP590      --CHHMA-ETTNALNENDRELRMMSRNLSFQLSVAMLFLEEAEERLTHVVRYGW--VPA
LmjF.36.5250  MSSEPLKAAAVTAVVASRECNLESL-EDVALAMQAAVQLVERQQGNINQLIAGGGDLHSGV
. . :   . * . : . . : * . : . : : : . . * : : * . : . : : :   *   .

TbBBP590      KVVSAALCGWLVLCCAGDLFHCSSGDTISTLVDRKELDDLRKLCGSQNAFDHGEVANVS
LmjF.36.5250  ESLNGRLCGVAKECDVTRPSLASLEETHGMLQ---RRMAQLE-----RHHV---DTIS
: . .   ***   * .   . * : * . *   . : : : * .   .   *   . : *

TbBBP590      RSECLLAIKRSVASLCADGAADPSLPLPTQLQKCVQSTVDEMYRLRDDADSALS ENVRLK
LmjF.36.5250  TVDCRA-----VTVL-----GAELR
: *   **   . . :

TbBBP590      RDLQMVVGGCETSLEYLGS--ACGSSSLCPALSPRSGVAAVSNLVMTARIASEKLSSLR
LmjF.36.5250  RLDKSTPADCLDGIAYVADALACRAEC-----
*   : . . . * . : : * . .   ** : . *

TbBBP590      VCEECNSVIERARRREERSHSRESSPASHRVEDGAEALLVERCKELVSELLVLRTRDRTS
LmjF.36.5250  -----DAAKAQLLAWRAEERIR
.   : : * .   * : : *

TbBBP590      TAAIDQRLMESELMVTSVRNAIENLR TSLDIDVPIGTPALDLVRTSGSCGEAVEGAISE-
LmjF.36.5250  EEASLR----TTTRLSTL---WQAFRTAL-----LPLLVD E PADRGCGCEVAVFSADPV
*   :   :   : : :   : : * : *   : *   * . *** . .   : .

TbBBP590      -----K
LmjF.36.5250  EFLASEAEARQAAQMIVDTAERLRRTHNDLQAVYAALRGHFGGDAATGVREMTATPHAPA

TbBBP590      RQVSATGPSLAELDTSIRTQLELCVLAVIRQLREKDRVISAAATLLGDALRSCSSAIRDE
LmjF.36.5250  RVLHNSGVTS LRFDDTVTATLHLVSTVVNAKCTSVK---DFIDHVVS DVLRDTSGSARV-
*   :   : * :   . : * : : : * . *   . *   :   .   .   .   : : * . *** . * : : *

TbBBP590      VTDSLLAVASG-TAHSLSLSL FSELQSVRASLKQSQHHT EELRTRLEGETTQLATLEAQ
LmjF.36.5250  ---SFAALRPGAFERLQGVLASLRALRESSDQHRDRA-----
* : * :   *   * . * : : : : : * * . *   : : :

TbBBP590      TESVATAERLLNEALQRQVAASQ TARLQRTLMDFAGSIAASTLYADDAGDRVDK-----
LmjF.36.5250  -----RRLPAFMDALVET---IHSHGGRVDVASTDA
* *   *   . : : . :   . .   * . ***

```

TbBBP590 LmjF.36.5250	TGDTV GASVS VVE-----ERWLDGI VRHVA QRMQ ESESLKGK LKLV TSE TGALAI ADATMPYRHSPAT TLEDCQAREQRAIVHGLQALLNQ E-----ERVRA LTAE ** :. : : :. :. : * : : * : : : * : *
TbBBP590 LmjF.36.5250	LESVRIRMTATEVAAREAE LNAQSVEVGNL TRIAEATARVELLRNAL SERSVELQKMHQR WQCVKM QYHQ--LSQE-----Q-----VAAEETVVELRRR-VQFKVQEDYKL--- :. * : : : :. :. : * : * : * : * : *
TbBBP590 LmjF.36.5250	VERHEKQVETLKA EISNDARSFDNERCSLLRAQSTLLAENTLKQQQLDEL RQKLWEGEKS ---EESLRELD SHLD-----QQARELSMR YFADQDA * :. :. * :. :. : * : * : : : :. : :
TbBBP590 LmjF.36.5250	LHEFLHNSRVSLVDATNKIRSL SANAVCGVLM LDEMERRSEILQSYI--LSFPTQEST I-----VRRFTLLRNQIHAAMRLPTRR-- : * : * : * : * : * : * : *
TbBBP590 LmjF.36.5250	LLSVCACIASTLSEREVQRLPDSVHRLNQ RVQKYKHL LGNPTKVP GPTRAVAVLALQKLA -----SASAMPRT PGRD TSTPRS----- . * :. * : *
TbBBP590 LmjF.36.5250	ETVELAESVSSTS FVSNLIAAVADFRTRYLI ANERIAKLEKEVNSGGTLMEMCDRRVWEM -----
TbBBP590 LmjF.36.5250	VKLLRQANSVGG LHSQS QEGPRGECQTDVVD FSVAQLDALQSLALECGIAAKYIHRHTEA -----
TbBBP590 LmjF.36.5250	NRKALETLRHVYDGESIFEACSHANEMMMKYRESEEDLRAQLLDVRSQLEACRHSIEESF -----
TbBBP590 LmjF.36.5250	SVFAETSTLND AKKVEFLQQQCAAAAELRRLKDEARVAQGLLASAGHDTHDGDICRNLA -----
TbBBP590 LmjF.36.5250	ELLKEEEKQRECVNLLQSEKKCARKQLQELGNKL RATEEMRNAELEE VQNRNDTLRTRLE -----
TbBBP590 LmjF.36.5250	EQSRQNNEWCQFVSTCAGVVGCPVENMQFEKSKNKILQTLRNVASSQSVCAEAGILCTTA -----
TbBBP590 LmjF.36.5250	QEEQLQCHREGANAAVARLDMPGSE RYRCQLLEFVGSVEEV LSELCP EKSESF LTTKDAP -----
TbBBP590 LmjF.36.5250	LELRVCQAAGQLCSAKDTVQANI QKLEAEIASAKKKNK AQKREMEDARKALHKELEVGGP -----
TbBBP590 LmjF.36.5250	AVASCALSTL TEGIRKKLRETLDENRKLGTALEELRIDNAQLERTLRCVEQQ LSETH TQL -----
TbBBP590 LmjF.36.5250	QCITESNGECRGQ NATLQERV DPLTEPT EEA SRKFRSPTQSPDIDADAAGLAQVHEVCFG -----
TbBBP590 LmjF.36.5250	RGPSIQHRSEEPREFEDNL RGNCRSDAAAVMDVG VAPKEQRSLLERLQQVTKERDVL TLE -----
TbBBP590 LmjF.36.5250	SEKLVERVQSVLYEHLNVQ QELLSYCSNANENELGESDTTVGDIVQLCVERLEHYLSETK -----
TbBBP590 LmjF.36.5250	TLREDATSAKLESARMAEEVAALRGRIKHLQDDNHVAERDSL RQNAEFESLSRQLQAFFN -----
TbBBP590 LmjF.36.5250	DALHMAVNDLRLRLPPVTD TSAIGALYGLQSLFDCMLS KCRDSIAFDEEKLNCRLNELKG -----

TbBBP590	KHAKDLKSLHRVFCAHVTITYTDDPNDEDSARSEGDGSSPVNLPTTDRLSGGLTPEHIVAA
LmjF.36.5250	-----
TbBBP590	GDALHHVHSNVCHIVWACIRHAERAPPPSLHTMGLPALTALLGEVVVASLERAEELLQAM
LmjF.36.5250	-----
TbBBP590	VAGSRALGILPVDPSLNQQHMEEWLECFITSEIQRVQRSSEGCCELLLNHLEILLQDHGISV
LmjF.36.5250	-----
TbBBP590	GELVSNRSQSSTLSNVGRSLQDNQLGSPRGPSSSVAVQDVMEFRSRALLGALRDLIAQFE
LmjF.36.5250	-----
TbBBP590	HRYSTLTSEWQALTDQNNRLKTSQQANEEEFARVQEQMQELRHAVRKKVEDDRKVEQSLR
LmjF.36.5250	-----
TbBBP590	DLDNHLDQMARELAMKYRADQEAISRQFAELRGSIRRVVIPITQEQLGTPSPKYDFVYPRE
LmjF.36.5250	-----

**Figure 9.7 (A) Sequence alignments of TbBBP590 against BLAST search top hit in *T. cruzi*, and (B) in *L. major*.**

P.falciiparum	-----MNI AIDEYGPFFVILREE-EKKRIK GIEAHKSNILA AKVVADI
G.lamblia	-----MAQRINGAQL LGVSDENGPNVYLMREGTEKKRARGLEAHKSNIT IARALANT
S.cerevisiae	MAARPQQPPMEMPDLSNAIVAQDEMGRPF IIVKDQGNKKRQHGLEAKKSHILAARSVASI
L.major	-----MSLAFDEYGRPFLLVKEQQQKERVSGVEAQRANILAALGVANV
T.brucei	-----MSLAFDEYGRPFLLVKEQANKERVGTGVEAQKANILAAISVSNV
T.cruzi	-----MSLAFDEYGRPFLLVKEQAQKERLTGVEAQKSNILAAISVSNV
P.tetraurelia	-----MSLAFDDYGRPYIILRDQGQKKRMKGLEAYKSNILA AKAVANT
D.discoideum	-----MSLVFDEYGNPFIVIRDQQA KERLRGIEAHRSHILA AKTISNI
Ch.reinhardtii	-----MSLAFDEFGRPF IILREQERKSRVKGIEAVKASIMAAKT VART
C.elegans	-----MAQSSAQLLFDESGQPFIVMREQENQKRITGVEAVKSHILAARAVANT
H.sapiens	-----MASMTLAFDEYGRPF IILKQDRKSRLMGLEALKSHIMAAKAVANT
C.owczarzaki	-----MSLAFDEYGRPF IILREQDTRQRTGSDAIKQHILA AKIAHT
	: * : * . * : : : : . * * : * : * * : :
P.falciiparum	LKSSIGPRGMDKIIIVSEDNNVTVTNDGATILEKIDVQHECAKLLVELSKSQDNEIGDGT T
G.lamblia	LRTSLGPRGMDKILVSGDREL TITNDGATILERMEVDNHIARLLVDLSKQDNEIGDGT T
S.cerevisiae	IKTSLGPRGLDKILISPGEITITNDGATILSQMDLDNEIAKLLVQLSKSQDDEIGDGT T
L.major	LKSSLGPRGMDKILITPDNEVVVTNDGATILDLMDIDNEVQQLMVLSKSDSEIGDGT T
T.brucei	LKTSLGPRGMDKILVTQDNEVVVTNDGATIMDLMDIDNEIGQLMVLSKSDSEIGDGT T
T.cruzi	LKTSLGPRGMDKILVTQDNDVVVTNDGATIVDLMDIDNEIGQLMVLSKSDSEIGDGT T
P.tetraurelia	LQSSLGPGKMDKMLVSPDGDVSVTNDGATIVEKMEIQHPTAKLLVELSQSDAEIGDGT T
D.discoideum	MKSSLGPGKMDKMMVSDGEVLVTNDGATILENMQVDNQIAKLMVQLSKSQDDEIGDGT T
Ch.reinhardtii	LRSSLGPGKMDKMLQSPDGDVTITNDGATILEMMEVENQIGKLMVELSKSQDHEIGDGT T
C.elegans	LRTSLGPRGLDKMLVSPDGDVTITNDGATIMEKMDVQHVVAKLMVELSKSQDHEIGDGT T
H.sapiens	MRTSLGPNGLDKMMVDKGDVTVTNDGATILSMDVDVHQIAKLMVELSKSQDDEIGDGT T
C.owczarzaki	LKTSLGPGKGLDKMMVSPDGAVTITNDGATILSQMDVQHPIAKLLVELSQSQDSEIGDGT T
	::*:*:*.*:*: : * : :*****. : : : : :. :*:*:*:*.** *****
P.falciiparum	GVVVIAGVLEEAYALIDKGIHPLRIADGFENACNIALKVIEDIALTVDIEEND--HKIL
G.lamblia	GVVVLASGLLEHAETLLDKGIHPRIISDGYDMACDVAIKHLENIAITTYFSADN--KEPL
S.cerevisiae	GVVVLASALLDQALELIQKGIHPKIANGFDEAAKLAISKLEETQDISASNDELIRDFL
L.major	GVVCFAGALLEQATALLDKGIHSSRISEGYEKACEMACKRLEEVEESISLEK---KEVL
T.brucei	GVVCLAGALLEQASGLLDKGIHSSRISEGFKEACEIACKRLEEIDTVPVSREE--YSYL
T.cruzi	GVVCFAGALLEQANGLLDKGIHSSRISEGFKEGCEVACKRIEEISETVPVSREE--YNYL
P.tetraurelia	GVVVLAGKLEQALTLLDKGIHPLKISDGFDRACDVAIKHLETISEEIDIQANE--HEAL
D.discoideum	GVVVLASGLLEQAEQLIEKGIHPRIYEGYETACKIATEHLKTIISDSIEFSKDN--IEPL
Ch.reinhardtii	GVVVLAGALLEHAELALDGMHPLRIAEGYEMACKVATENLDNIHTRFDFSKEN--IEPL
C.elegans	GVVVLAGALLEEAELIDRGIHPRIKIADGFDLACKKALETLDISIDKFPVE--NRERL
H.sapiens	GVVVLAGALLEEAELQLDRGIHPRIADGYEQAAARVAIEHLDKISDSVLVDIKD--TEPL
C.owczarzaki	GVVVLASGLLEQAEHLIDRGIHPRIADGFEMACAVAVKQLEAVSDIIPVGKDE--KESL
	*** :*. ** :*. * : : * : * : : : . . * . . .



P.falci parum	KKVAKTSLSSKIVSSKKDLNSNIVDAVLSVADMKRKDVRFDLIKIEGKTGGLLESTLI
G.lamblia	IKAAMTSLGSKVINRYHRTMAEVAVDAVLRVADLQRNDVDFELIKVRSKIGGKLEDIKII
S.cerevisiae	LRAAKTSLGSKIVSKDHDREAFEMAVEAVINVMMDKDRKDVDFDLIKMQGRVGGSSISDSKLI
L.major	LQTARSTLNSKVVNRDRDLAQICVDAVLAVADMERRDVNMDLIKVEGKVGGRLEETCLV
T.brucei	LQTARSTLNSKVVNRDRDLAKICVDAVLAVADMERRDVNLDLIKMEGKVGGCLEETCLV
T.cruzi	LQTARSTLNSKVVNRERDRDLAKICVDAVLAVADMERRDVNLDLIKVEGKVGGCLEETCLV
P.tetraurelia	IQAAC TALGSKVVS KRKREL GKIAVTAVLDVADLARKDVNLDLIKIQTKTGGSVEDTRLI
D.discoideum	IKTAMTCLGSKIVNRFHRQMS EIAVKAVISVADLERKDVNLENIKLEGKEGGKLEDTLV
Ch.reinhardtii	VKTCMTTLSSKVVGR LKRPMAEICVKAVLAVADLERRDVNLDLIKVEGKVGGMEDTRLV
C.elegans	VETAQTS LGSKIVNRS LRQFAEIAVDAVLSVADIESKDVNFEMIKMEGKVGGRLEETILV
H.sapiens	IQTAKTTLGSKVVNSCHRM AEIAVNAVLTVADMERRDVDFELIKVEGKVGGRLEETKLI
C.owczarzaki	VQTAMTTLGSKIIINKCHRM AEIAVDAVLAVADLETRDVNFELIKMEGKVGQQLQDTQLV
	... : * . * * : : : : * * : * * . * * : : * * : : * * : : *
P.falci parum	KGIVLNKELSHSQMIKEV-----RNAKIAILTCFPEPPKPKIKHKNITNVDAYRDLQA
G.lamblia	NGFLIDQEMSHPGMRKSI-----NHAKIAVLTCTFPEPPKPKTTYKIDVDSVEKYNTLYK
S.cerevisiae	NGVLDKDFSH PQMPKCVLPKEGSDGVKLA ILTCFPEPPKPKTKHKL DISSVEEYQKLQT
L.major	DGIVLDKDFSH PQMPKEL-----TNPKMALTCAFEPKPKTKHTLQINS AEHMRELHE
T.brucei	NGIVLDKDFSH PQMPKVL-----KNPKIAILTCFPEPPKPKTKHTVHISAEHMKEIHE
T.cruzi	NGIVLDKDFSH PQMPKQL-----KNPKIAILTCFPEPPKPKTKHTVQIANAEQMS EIIHA
P.tetraurelia	SGILIDKDM SH PQMPKEV-----KEAKICLLTCFPEPPKPKTKHNINISSAEDYKKLYQ
D.discoideum	KGIIIDKGISH PQMPKII-----KDAKICLLTCFPEPPKPKTKNSIEITKAEDFKVLGE
Ch.reinhardtii	EGIVIDKEF SH PQMPKEL-----KDVKVAILTCFPEPPKPKTKHKVDIDTVEKF EALRD
C.elegans	KGIVIDKTM SH PQMPKEL-----KNAKVAILTCFPEPPKPKTKHKLDITSTEDFKALRD
H.sapiens	KGIVIDKDFSH PQMPKKV-----EDAKIAILTCFPEPPKPKTKHKLDVTSVEDYKALQK
C.owczarzaki	RGVVIDKDFSH PQMPKEL-----RDAKIAILTCFPEPPKPKTKHKLDIASVEDYKLSQS
	* . : : : : * * * * : * . : : * * * * * * * * . : : : : * : :
P.falci parum	IEKKYFYDMVASLKKAGANFVICQWGFDD EANYLL LKENIPAIRWVGGVEMELIAIATGG
G.lamblia	EEQQFFINMIEHLKGAGANFVACQWGFDD EANHLLMQNDINAVRWVKGEDIEAIALATGA
S.cerevisiae	YEQDKFKEMID DVKKAGADVVICQWGFDD EANHLLQNDLP AVRWVGGQELEHIAISTNG
L.major	QEQEYFRSEVQRCKDAGADLVICQWGFDD EANYLLYRNQLPAVRWVGGVELEMIAIATAG
T.brucei	QEQEYFRNQVKLCKQAGADLVICQWGFDD EANYLLYRNQLPAVRWVGGVELEMIAIATGG
T.cruzi	QEQEYFRKQVQLCKDAGAE LVICQWGFDD EANYLLFRNDLP AVRWVGGVELEMIAIATGG
P.tetraurelia	QEQQYFVDMVKLVKDSGANIALCQWGFDD EANHLLQNELPAVRWVSGTDVELIAMATGA
D.discoideum	IEQKYFTDMVEKVKATGANLVICQWGFDD EANHLLQNNLP AVRWVGGGLDLEKIAMATGG
Ch.reinhardtii	AEKSYFTDMVQRCKDSGAELVICQWGFDD EANHLLMHKNLP AVRWVGGVEIELLAMATGA
C.elegans	YERETFTMIRQVKESGATLAICQWGFDD EANHLLQANDLP AVRWVGGPEIELLATATNA
H.sapiens	YEKEKFEEMIQQIKETGANLAICQWGFDD EANHLLQNNLP AVRWVGGPEIELIAIATGG
C.owczarzaki	YEHKIFVEMI ERVKQSGANFVVCQWGFDD EANHLLQNSLP AVRWVGGPEIELVATATNG
	* . : * : * : * * . . * * * * * * * * * * : : * : * * * : * : * : *
P.falci parum	KIIPRFEDIDESKL GKADLIREISHGTVNNPMVYIEGCSNT-----KAITILLRGG
G.lamblia	RIIIPRFEDISSDKLGYAEVVREVS LGTSSDKVMVFEGLQPCQT TADSVPIMISTVFIRGG
S.cerevisiae	RIVPRFQDLSKDKLGTCSRIYEQEF GTTKDRMLIEQSKET-----KVTTCFVRGS
L.major	RIVPRFEDLTAEKLGTCGRVREVVGFGTTKDRMIFVENC PNS-----KAVTVFIRGG
T.brucei	RIIPRFEDLQSAKLGTCGTVREVVGFGTTKDRMIFIEECPSS-----KAVTIFIRGG
T.cruzi	RIIIPRFEDISGSKLGTGCTVREVVGFGTTKDRMIFIEDCPNS-----RAVTILIRGG
P.tetraurelia	RIIPRFQEITPEKLGKAGSVKEVQFGTSNERMLVIE NCQCT-----KAVTILIRGG
D.discoideum	RIVARFEDVSADKLGRAGLVREVVGFGTTQDRYLSIEDCPNT-----NAVTFIVRGG
Ch.reinhardtii	RIVPRFQELSADKLGHAGSVKEVGF GTTKDKMLVIEGCPHI-----KAVTIFVRGG
C.elegans	RIVPRFSELSKEKLG TAGLVREITFGAAKDRMLSIEQCPNN-----KAVTIFVRGG
H.sapiens	RIVPRFSELTAELKGFAGLVQEISFGTTTKDKMLVIEQCKNS-----RAVTIFIRGG
C.owczarzaki	RIVPRFEELSPEKLGHAGVREISFGTTTKDRMLFIEDCKNT-----RAVTIVRGG
	: * : * * : : * * * . : * * : . : * * : : * * : : * * : : *
P.falci parum	NQMMIEECERSVHDALCSVRNLIR-----NRILPGGGSTEIYAALEIEKVADKCKGIEQYAIRA
G.lamblia	NNMIEEARRSLHDAMCVVRNLIR-----OPRVIYGGGSAEISCSLAVAEAAEHVGTVEQYALRA
S.cerevisiae	NKMIVDEAERALHDSL CVVRNLVK-----SRVVYGGGAAEVTMSLAVSEADKQRGIDQYAFRG
L.major	NKMMIEEAKRSLHDAICMVVRNLIR-----ONRIVYGGGSAELAASSAVLAHADAVSTVDQYAIRA
T.brucei	NKMMIEEAKRSLHDAICMVVRNLIR-----ONRIVYGGGSAEIAASMAVLNYADTVSTVDQYAIRR
T.cruzi	NKMMVEEAKRSLHDAICMVVRNLIR-----ONRVVYGGGSAEIAASLAVLNYADSISTVDQYAIRR
P.tetraurelia	SQMIVDEAKRSIHDAICVIRNLIK-----ONRIVYGGGSAELACSIQVLQYADEVSSVEQYAIRS
D.discoideum	NKMIVEEAKRSIHDAICVTRNLIR-----ONRIVYGGGSS EISGLKISAMADDIASIEQYAVRA
Ch.reinhardtii	NRMVLDEIKRSLHDAICVARNLVR-----NAIVYGGGAAEISCALAVEEAADKVTGVEHYAMRT
C.elegans	NKMIIDEAKRALHDAICVIRNLVR-----OSRIVYGGGSAELAAAIQVAKEADRIDGIEQYAFRA
H.sapiens	NKMIIEEAKRSLHDAICVIRNLIR-----ONRVVYGGGAAEISCALAVSQAEDKCP TLEQYAMRA
C.owczarzaki	NQMIIEEAKRSLHDAICVVRNLVK-----ONRVVYGGGAPEITCSLAVYEAAADKVATLEQYAMRA
	. * : : : * . * : * * : * * * : : : * * : * : * : * : : * * : * * *
P.falci parum	FGNALLSIPINLCNNMGLNSIDIISEIKTKIIQDKTKNLGIDCLNYKVGDMIERGIFETFL
G.lamblia	FSEALDIIPLTLARNSSIQPIETLAMVKAQKKTGTHHLGIDCMRRNTIDMKEQNVFETFL
S.cerevisiae	FAQALDTIPMTLAENSLDPIGTLSTLKSQQLKEKISNIGVDCLGYSGNDMKELFVDDPF
L.major	FADALESIPINLALNSGLDPIRALSDARKAQIQGNPNPYAGIDCMDRGTIDMKQQEVFETFL

T.brucei	FADALESIPINLALNSGLEPIKCLSRARIEQVEGKNPYAGVDCMDSGTLDMMKKQQVFETL
T.cruzi	FADALESIPINLALNSGLDPIKALSLARIVQVEGKNPYAGVDCMDCGTIDMKQQQVFETL
P.tetraurelia	FADALEGIPCCCLADNCGLNPITSVAAQAKARQLQESNPRIIDCMELGTTDMKEQRIYETL
D.discoideum	FADALDAIPLALAENSGLPISIESLSTVKAMQIKEKNPRLGIDCNHRDTNDMKAQHVFDTL
Ch.reinhardtii	FADALQAVPLALAENSGLPPIESLTAVKKRQLEEKNPYLIDCNDTGTNDMREQNVFETV
C.elegans	FADALESIPMALAENSGLPIDALSDLKAKQIETGKSSLGIDAVFAGTNDMKEQKVIETL
H.sapiens	FADALEVIPMALENSGMNPIQTMTEVRARQVKEMNPALGIDCLHKGTDNDMKQQHVIETL
C.owczarzaki	FGDALEATPIALAENSGLPPIETLAAVKSRQVKEKNPRLGIDCNMTGNADMRDQHVETL
	*.:** * *. * .: * :: : : *:*. ** : : .
P.falciparum	NSKYNQFSLATQVVKMILKIDDVAPNDFN----
G.lamblia	SSKVQQLLATQVVRMILKIDEVITDEV-----
S.cerevisiae	IGKKQQILLATQLCRMILKIDNVISGKDEY---
L.major	SGKCSQLRLATQVVKMILKIDDVITKPDDEEQQ
T.brucei	QGKCSQLRLATQVVKMILKVDDVITKPEEEE--
T.cruzi	QGKCSQLRLATQVVKMILKVDDVILTHVEEQ--
P.tetraurelia	SSKKQQIQLATQVVKMILKIDDVAPDDY-----
D.discoideum	PGKTQQFLANQVVKMILKIDDIKMGPGADE--
Ch.reinhardtii	MGKKQQFLATQVCKMILKIDDVVKPSDYE----
C.elegans	LSKREQISLATQVVRMILKIDDVVPDDERMGY-
H.sapiens	IGKKQQISLATQVMVRMILKIDDIKPGESEEE---
C.owczarzaki	VGKKQQFQLATQLVKMILKIDDIPTPGGAQ----
	. * .*: **. *: :****:***:

**Figure 9.8 Sequence alignments of the TCP-1-ε protein in different organisms.** GroEL-like equatorial domains that locate at both C and N terminus are in **red bracket**; (2) GroEL-like apical domain is located at the central of the sequences in **orange bracket**; (3) TCP-1-like chaperonin intermediate domain, which can be found between the GroEL-like equatorial and apical domain shows in **blue bracket**.

## Appendix III

The RNAi constructs of all eleven selected putative *Tb*RP2-interacting /nearby neighbour proteins analysed by Sanger Sequencing. In all the sequences analyses restriction sites BamHI (GGATCC) and XhoI (CTCGAG) (Sall; GTCGAC for *Tb*KMP11) are highlighted in different colours, RNAi targeted region is highlighted in bold.

### 9.9 Arl3L RNAi sequence analysis (*Tb*927.6.3650)

GAATTCAGTAGTGATTGCGGGATCCTGAGCCGCGTGACTGATAGTGGGCCTTGACAATGCT  
**GGAAAGACTACCGTACTGAATGCATTGGGAGAAGACGAGGTGCCTGTGGAGGGAAAAGTTTC**  
**CCACGCCGCACCTGAGGGACCAACGCAAGGGTTTAATATCAAAACACTCACACGTGGGAATA**  
**AACGGGCTAAGTTATGTGATCTCGGTGGCCAGCGGGCGCTGCGAGACTATTGGCAGGATTAC**  
**TACAGCAACACGGACTGCATTATGTACGTGGTGGATTCTTCTGACCACCGCCGACTCGAAGA**  
**ATCTCACGCTGCATTTGTGGATGTTCTAAAGGGCATTGAAGGTGCTCCGGTTTTGGTGTTTG**  
**CGAATAAACAAGATCTTGCAACCGCAAAAGATGCTCAGGCGATCGCAGAATGCCTTCATTG**  
**CATGACTTCCGAGACCGCAAGTGGCACATCCAAGGCTGTAGCGCCAAAACGGGAGCAGGACT**  
**GGAGGAAGGCGTGACATGGATACTTAGCACTTGTGCCCCGTAATGGGAAGCTGATGTCTGTA**  
**AGGCGCTGAAAGCAAATAA****CTCGAG**CGCAATCGAATTC

#### ➤ Translated protein sequence:

ggatcctgagccgctgtactgatagtgggccttgacaatgctggaaagactaccgtactg  
D P E P R V L I V G L D N A G K T T V L  
aatgcattgggagaagacgaggtgcctgtggagggaaaagtttccacgccgcacctgag  
N A L G E D E V P V E G K V S H A A P E  
ggaccaacgcaagggtttaatatcaaaacactcacacgtgggaataaacgggctaagtta  
G P T Q G F N I K T L T R G N K R A K L  
tgtgatctcgggtggccagcgggctgagactattggcaggattactacagcaacacg  
C D L G G Q R A L R D Y W Q D Y Y S N T  
gactgcattatgtacgtggtggattcttctgaccaccgccgactcgaagaatctcacgct  
D C I M Y V V D S S D H R R L E E S H A  
gcatttgtggatgttctaaagggcattgaaggtgctccggttttggtgtttgcaataaaa  
A F V D V L K G I E G A P V L V F A N K  
caagatcttgcaaccgcaaaagatgctcaggcgatcgacagaatgccttcatttgcacgac  
Q D L A T A K D A Q A I A E C L H L H D  
ttccgagaccgcaagtggcacatccaaggctgtagcgccaaaacgggagcaggactggag  
F R D R K W H I Q G C S A K T G A G L E  
gaaggcgtgacatggatacttagcacttgtgccccgtaatgggaagctgatgtctgtaag  
E G V T W I L S T C A P - W E A D V C K  
gcgctgaaagcaaataa**ctcgag**  
A L K A N N S

### ➤ TritrDB BLAST alignments:

```
> Tb927.6.3650:mRNA-p1 | transcript=Tb927.6.3650:mRNA | gene=Tb927.6.3650
| organism=Trypanosoma_brucei_brucei_TREU927 | gene_product=ADP-ribosylation
factor, putative | transcript_product=ADP-ribosylation
factor, putative | location=Tb927_06_v5.1:1087819-1088370 (-)
| protein_length=183 | sequence_SO=chromosome
| SO=protein_coding
Length=183
```

```
Score = 353 bits (906), Expect = 3e-125, Method: Compositional matrix
adjust.
Identities = 169/170 (99%), Positives = 169/170 (99%), Gaps = 0/170 (0%)
```

```
Query 1 EPRVLIVGLDNAGKTTVLNALGEDEVPEVGKVSHAAPEGPTQGFNIKTLTRGNKRAKLCD 60
EPRVLIVGLDNAGKTTVLNALGEDEVPEVGKVSHAAPEGPTQGFNIKTLTRGNKRAKLCD
Sbjct 14 EPRVLIVGLDNAGKTTVLNALGEDEVPEVGKVSHAAPEGPTQGFNIKTLTRGNKRAKLCD 73

Query 61 LGGQRALRDYWDYYSNTDCIMYVVDSSDHRRLSESHAAAFVDVLKGIEGAPVLVFANKQD 120
LGGQRALRDYWDYYSNTDCIMYVVDSSDHRRLSESHAAAFVDVLKGIEGAPVLVFANKQD
Sbjct 74 LGGQRALRDYWDYYSNTDCIMYVVDSSDHRRLSESHAAAFVDVLKGIEGAPVLVFANKQD 133

Query 121 LATAKDAQAIAECLHLHDFRDRKWHIQGCSAKTGAGLEEGVTWILSTCAP 170
LATAKDAQAIAECLHLHDFRDRKWHIQGCSAKTGAGLEEGV WILSTCAP
Sbjct 134 LATAKDAQAIAECLHLHDFRDRKWHIQGCSAKTGAGLEEGVAWILSTCAP 183
```

## 9.10 *TbCAP51/TbCAP51V* RNAi sequence analysis (*Tb927.7.2640; Tb927.7.2650*)

```
GAATTCAGTAGTGATTGCGGATCCGGAAGAAACCGAGGAGCAGATTATTATTCGGC
TGCAGAAGGAGCTGAGGGAATTCGAAGCATATGAGGAGGAGATGCAGCCAAACCTGA
ACATGTACGAAGAGGTCATCGAACTCCGACGAGAAGCTGAGCAACTGGAGGATGACA
GGAAGCGCGCATTTGCTTCAGATTGATCTTGGTGATATATCCTCTTACTACGACGGTG
ATGTGGAACGGGAGGTGGAGGAGCTGGAGGCAGCAGTGGAGAATTCATATTTACAAG
CCTGCTGGACGAAAGAGCGCCAACGATACATCGCAAGGAACGAGACGGACCTTCTTG
AGTGGTGGCGTGTGGATATCATGAAGTTGGGATTTCTGTTTGAGGGAGGCCAAAAGGT
CATTTGATTCAAGTAATCGTGTGTCCTCAAGACTTATATCGTGCCACGAGGTGAAGG
TGAACGCTACCAACGACTCTCGTGACAAGGCAAAGCTGAAGTGCAGCAAGAGATCG
CTGCGGAGTTGAACGAGCTGCGTCAGCGCGGGAACGTCTGTTGGTTATAGACAAGG
AACAAACGTTTCCATTTACGCCGTGGTACTCACGTGAAGGCTCGAGCGCAATCGAATT
C
```

### ➤ Translated protein sequence:

```
ggatccggaagaaacccgaggagcagattattatttcggctgcagaaggagctgaggggaattc
D P E E T E E Q I I I R L Q K E L R E F
gaagcatatgaggaggagatgcagccaaacctgaacatgtacgaagaggtcatcgaactc
E A Y E E E M Q P N L N M Y E E V I E L
cgacgagaagctgagcaactggaggatgacaggaagcgcgcatctgcttcagattgatctt
R R E A E Q L E D D R K R A L L Q I D L
ggtgatatatcctcttactacgacgggtgatgtggaacgggaggtggaggagctggaggca
G D I S S Y Y D G D V E R E V E E L E A
gcagtggagaattcatatttacaagcctgctggacgaaagagcgccaacgatacatcgca
A V E N S Y L Q A C W T K E R Q R Y I A
```

aggaacgagacggaccttcctgagtgggtggcgtgtggatatcatgaagttgggatttcgt  
 R N E T D L P E W W R V D I M K L G F R  
 ttgagggaggccaaaaggtcatttgattcaagtaatcgtgttgcccaagacttatatcgt  
 L R E A K R S F D S S N R V A Q D L Y R  
 gcccacgaggtgaaggtgaacgctaccaccgactctcgtgacaaggcaaaagctgaactg  
 A H E V K V N A T T D S R D K A K A E L  
 cggcaagagatcgctgcgaggttgaacgagctgctcagcgcggggaacgtctgttggtt  
 R Q E I A A E L N E L R Q R R E R L L V  
 atagacaaggaacaacgtttccatttacgccgtggtactcacgtgaagg~~ctcgag~~  
 I D K E Q R F H L R R G T H V K A R

### ➤ TritrDB BLAST alignments:

> Tb927.7.2650:mRNA-p1 | transcript=Tb927.7.2650:mRNA | gene=Tb927.7.2650  
 | organism=Trypanosoma\_brucei\_brucei\_TREU927 | gene\_product=hypothetical  
 protein, conserved | transcript\_product=hypothetical  
 protein, conserved | location=Tb927\_07\_v5.1:680735-682348 (-)  
 | protein\_length=537 | sequence\_SO=chromosome  
 | SO=protein\_coding  
 Length=537

Score = 362 bits (930), Expect = 1e-123, Method: Compositional matrix  
 adjust.  
 Identities = 187/189 (99%), Positives = 189/189 (100%), Gaps = 0/189 (0%)

Query	8	QIIIRLQKELREFEAYEEEMQPNLNMYEEVIELRREAEQLEDDRKRALQLIDLGDISSYY	67
		QIIIRL++ELREFEAYEEEMQPNLNMYEEVIELRREAEQLEDDRKRALQLIDLGDISSYY	
Sbjct	235	QIIIRLREELREFEAYEEEMQPNLNMYEEVIELRREAEQLEDDRKRALQLIDLGDISSYY	294
Query	68	DGDVEREVEELEAAVENSYLQACWTKERQRYIARNETDLPEWWRVDIMKLGFRLEAKRS	127
		DGDVEREVEELEAAVENSYLQACWTKERQRYIARNETDLPEWWRVDIMKLGFRLEAKRS	
Sbjct	295	DGDVEREVEELEAAVENSYLQACWTKERQRYIARNETDLPEWWRVDIMKLGFRLEAKRS	354
Query	128	FDSSNRVAQDLRYAHEVKVNATTDSDRKAKAELRQEIAAELNELRQRRLRLVIDKEQRF	187
		FDSSNRVAQDLRYAHEVKVNATTDSDRKAKAELRQEIAAELNELRQRRLRLVIDKEQRF	
Sbjct	355	FDSSNRVAQDLRYAHEVKVNATTDSDRKAKAELRQEIAAELNELRQRRLRLVIDKEQRF	414
Query	188	HLRRGTHVK	196
		HLRRGTHVK	
Sbjct	415	HLRRGTHVK	423

> Tb927.7.2640:mRNA-p1 | transcript=Tb927.7.2640:mRNA | gene=Tb927.7.2640  
 | organism=Trypanosoma\_brucei\_brucei\_TREU927 | gene\_product=cytoskeleton  
 associated protein, putative | transcript\_product=cytoskeleton  
 associated protein, putative | location=Tb927\_07\_v5.1:676956-678317 (-)  
 | protein\_length=453 |  
 sequence\_SO=chromosome | SO=protein\_coding  
 Length=453

Score = 299 bits (766), Expect = 4e-100, Method: Compositional matrix  
 adjust.  
 Identities = 156/189 (83%), Positives = 171/189 (90%), Gaps = 0/189 (0%)

Query	8	QIIIRLQKELREFEAYEEEMQPNLNMYEEVIELRREAEQLEDDRKRALQLIDLGDISSYY	67
		QIIIRLQKELREFEAYE+EMQPNLN+YE++IELRREAEQLEDDRKRALQLIDLGDISSYY	
Sbjct	151	QIIIRLQKELREFEAYEKEMQPNLNIYEDIELRREAEQLEDDRKRALQLIDLGDISSYY	210
Query	68	DGDVEREVEELEAAVENSYLQACWTKERQRYIARNETDLPEWWRVDIMKLGFRLEAKRS	127
		DGDVEREVEELEAAVENSYLQACWTKERQRYIARNETDLPEWWRV++ L RL++A+ S	
Sbjct	211	DGDVEREVEELEAAVENSYLQACWTKERQRYIARNETDLPEWWRVNVKDLPSRLKDARAS	270
Query	128	FDSSNRVAQDLRYAHEVKVNATTDSDRKAKAELRQEIAAELNELRQRRLRLVIDKEQRF	187

```

Sbjct 271 FDSSNRVAQDLYRAHEVKVNATTDSDRKAKAE+ + AE+ + RERL + K+Q F
          FDSSNRVAQDLYRAHEVKVNATTDSDRKAKAEIFNGMKAIEATKGERERLRQLQKDQLF 330

Query 188 HLRRGTHVK 196
          HLRRGTHVK
Sbjct 331 HLRRGTHVK 339

```

## 9.11 *Tb*MARP RNAi sequence alignments (*Tb*927.10.10360)

GAATTCGATT**GGATCC**GGCCACTTTAGGACAACCTGCAAAGGATTTCGTATTGTCGCTA  
TGATTCTTCTTGTATAAGAGTTCGTGTAGGGAGTCTTCTTTGGTGATGAATCTGT  
TATATCTAAGAGGCATAGCTTTACTGCTTCATCTGTAGCTGATGCTATTGTTGTTAA  
GAAGGTCGTGAAAGGTAAAAGTGATGGTGGTGTGACATCAAAGAAGGTTGATGTCTC  
CACTGATATGAAGTGTGTCGCAAGTGGGGAGAGTAATGGAGGTGCTGCTGCGGTTGT  
TGCAGCTACTACTGAGTGTGCTGCTGCGGCTGTGGGCGATCAACCAACTGGTAACGG  
AGTACGTGTGTCAAAGTGGAAAAAGGCCCTAATGTTGGTTATTCCCTGCCCTTGTC  
TGTAGATGCTGATATGTACGTGAGTACCGCGCATAGAGATTTCAAGGCACACGGTGC  
CAGCAAACCTTATATGCCAAAGGCTGCTCCGACTGTTAAACAGTCAAGCATTTCTAT  
GCAAGGTGTAAGCAGCG**CTCGAG**AATCACTAGTGAATTC

### ➤ Translated protein sequence:

ggcgattgggcccgcacgtcgcatgctccccggccgcatggcgccgcgggaattcgatt**gga**  
R L G P T S H A P G R H G G R G N S I G  
**tcc**ggccacttttaggacaactgcaaaggattcgtattgtcgctatgattcttcttggtat  
S G H F R T T A K D S Y C R Y D S S C Y  
aagagttcgtgtagggagtccttcctttggtgatgaatctgttatatctaagagggcatagc  
K S S C R E S S F G D E S V I S K R H S  
tttactgcttcacatctgtagctgatgctattggtgttaagaaggctcgtgaaaggtaaaagt  
F T A S S V A D A I V V K K V V K G K S  
gatgggtggtgtgacatcaaagaagggtgatgtctccactgatatgaagtgtgtcgcaagt  
D G G V T S K K V D V S T D M K C V A S  
ggggagagtaatggaggtgctgctgcggttggtgcagctactactgagtgtgctgctgcg  
G E S N G G A A A V V A A T T E C A A A  
gctgtggggcgatcaaccaactggtaacggagtacgtgtgtcaaagtggaaaaaggcccct  
A V G D Q P T G N G V R V S K W K K A P  
aatgttggttattcctgccccttgctcatgtagatgctgatatgtacgtgagtaccgcgcat  
N V G Y S C P C H V D A D M Y V S T A H  
agagattttcaaggcacacggtgccagcaaaccttatatgccaaaggctgctccgactgtt  
R D F K A H G A S K P Y M P K A A P T V  
aaacagtcgaagcatttctatgcaagggtgtaagcagcg**ctcgaga**aatcactagtgaattcg  
K Q S S I S M Q G V S S A R E S L V N S

### ➤ TritrDB BLAST alignments:

```

> Tb927.10.10360:mRNA-p1 | transcript=Tb927.10.10360:mRNA | gene=Tb927.10.10360
| organism=Trypanosoma_brucei_brucei_TREU927
| gene_product=Microtubule-associated repetitive protein |
transcript_product=Microtubule-associated repetitive protein
| location=Tb927_10_v5.1:2548906-2558502(-) | protein_length=3198
| sequence_SO=chromosome | SO=protein_coding

```

Length=3198

Score = 300 bits (769), Expect = 9e-94, Method: Compositional matrix adjust.  
Identities = 166/171 (97%), Positives = 167/171 (98%), Gaps = 0/171 (0%)

```
Query 1 GHFRTTAKDSYCRYDSSCYKSSCRESSFGDESVISKRHSFTASSEADAIIVVKVVKGKSD 60
Sbjct 2979 GHFRTTAKDSYCRYDSSCYKSSCRESSFGDESVMRHSFTASSEADAIIVVKVVKGKSD 3038

Query 61 GGVTSKKVDVSTDMKCVASGESNNGAAAVVAATTECAAAVGDQPTGNGVRVSKWKNAPN 120
Sbjct 3039 GGVTSKKVDVFTDMKCVASGESNNGAAAVVAATTECAAAVGDQPTGNGVRVSKWKKAPN 3098

Query 121 VGYSCPCHVDADMYVSTAHRDFKAHGASKPYMPKAAPTVKQSSISMQGVSS 171
Sbjct 3099 VGYSCPCHVDADMYVSTAHRDFKAHGASKPYMPKAAPTVKQSSISMQGVSS 3149
```

## 9.12 *Tb*KMP11 RNAi sequence analysis (*Tb*927.9.13920)

GAATTCAGTAGTGATTGCG**GGATCC**AGATCGTGCTTTCAAGAGTAAACTAACTTTAAGA  
TACTTTTTTCAAGAAACATGGCCACCACATACGAAGAATTTGCTGCGAAGCTCGACCGCCT  
CGATGCCGAATTCGCCAAGAAGATGGAGGAGCAGAACAAGCGATTCTTCGCTGACAAGCC  
TGATGAGGCTACGCTGTCCCCTGAGATGAAAGAGCACTATGAAAAGTTTCGAAAAATGAT  
CCAGGAGCACACGGACAAGTTCAACAAGAAGATGCGCGAGCACTCAGAGCACTTCAAGGC  
CAAGTTTTCGCGAAGCTCCTCGAGCAGCAGAAGAATGCCAGTTCCCCGGAAAAATGATTGCA  
AATTTACTACACCTCTCATTTTCTTTCTATGTTAGAGTCCCTAGGGATAATACTATAAA  
AATCTTTTAAAAGATGTAAATACCACCTTTTCTAGCGCAATGC**GTCGAC**CGCAATCGAATT  
C

### ➤ Translated protein sequence:

gcgaattcactagtgattgcg**ggatcc**agatcgtgcttttcaagagtaaactaaacttttaa  
A N S L V I A G S R S C F Q E - T K L -  
gatacttttttcaagaaacatggccaccacatacgaagaatttgctgccaagctcgaccgc  
D T F S R N M A T T Y E E F A A K L D R  
ctcgatgccgaattcgccaagaagatggaggagcagaacaagcgattcttcgctgacaag  
L D A E F A K K M E E Q N K R F F A D K  
cctgatgaggctacgctgtcccctgagatgaaagagcactatgaaaagtttcgaaaaaatg  
P D E A T L S P E M K E H Y E K F E K M  
atccaggagcacacggacaagttcaacaagaagatgcgcgagcactcagagcacttcaag  
I Q E H T D K F N K K M R E H S E H F K  
gccaaagtttgcggaactcctcgagcagcagaagaatgccagttccccggaaaaatgattg  
A K F A E L L E Q Q K N A Q F P G K - L  
caaatttactacacctctcattttctttcctatgttagagtccttagggataatactata  
Q I Y Y T S H F L S Y V R V P R D N T I  
aaaatctttttaaagatgtaaataccacttttctagcgcaatgc**gtcgacc**gcaatcgaa  
K I F - K M - I P L F - R N A S T A I E

### ➤ TritrDB BLAST alignments:

> *Tb*927.9.13820:mRNA-p1 | transcript=*Tb*927.9.13820:mRNA | gene=*Tb*927.9.13820  
| organism=*Trypanosoma brucei brucei* TREU927 |  
gene\_product=kinetoplastid membrane protein 11-3 |  
transcript\_product=kinetoplastid

membrane protein 11-3 | location=Tb927\_09\_v5.1:2179681-2179959 (-)  
 | protein\_length=92 | sequence\_SO=chromosome  
 | SO=protein\_coding  
 Length=92

Score = 184 bits (468), Expect = 1e-60, Method: Compositional matrix adjust.  
 Identities = 92/92 (100%), Positives = 92/92 (100%), Gaps = 0/92 (0%)

```
Query 25 MATTYEEFAAKLDRLDAEFAKKMEEQNKRFFADKPDEATLSPEMKEHYEKFEKMIQEHTD 84
          MATTYEEFAAKLDRLDAEFAKKMEEQNKRFFADKPDEATLSPEMKEHYEKFEKMIQEHTD
Sbjct 1 MATTYEEFAAKLDRLDAEFAKKMEEQNKRFFADKPDEATLSPEMKEHYEKFEKMIQEHTD 60

Query 85 KFNKKMREHSEHFKAFAELLEQQKNAQFPGK 116
          KFNKKMREHSEHFKAFAELLEQQKNAQFPGK
Sbjct 61 KFNKKMREHSEHFKAFAELLEQQKNAQFPGK 92
```

## 9.13 *Tb*HSP70RNAi sequence analysis (*Tb*927.11.11330)

GAATTCGATTTATGGATCCATCGAGATCGATGCACTCTTCGAAAACATTGACTTCCAGGC  
 GACCATTAACCCGTGCTCGTTTCGAGGAAGTGTGCGGTGACCTGTTCCGTGGTACATTGCA  
 ACCCGTAGAGCGTGTGCTGCAAGGATGCCAAGATGGACAAACGCGCCGTGCACGATGTGGT  
 CCTGGTTGGTGGTTCCACCCGTATCCCCAAGGTGATGCAGCTTGTGTCTGACTTTTTTCGT  
 GGCAAGGAGCTTAACAAGAGCATCAACCCGACGAGGCAGTGGCCTACGGCGCTGCTGTA  
 CAAGCTTTCATCTTGACCGGTGGTAAGAGCAAGCAAACGGAAGGTCTCCTCCTTCTGGAT  
 GTCGCCCCGCTTACGCTCGGCATCGAGACGGCAGGTGGCGTGATGACGGCACTGATCAAG  
 CGCAACACGACGATCCCCACCAAAAAGAGCCAAATCTTCTCCACATACTCAGACAATCAG  
 CCCGGTGTGCACATTAGGTCTTTGAGGGTGAGCGTACCATGACGAAGGACTGCCATCTC  
 TTGGGCACCTTCGATCTCTCTGGAATCCCCCGGCACCCCGCGGCGTGCCACAGATAGAA  
 GTGACCTTCGACCTAGACGCTAATGGTATTCTCGAGCGCAATCACTAGTGAATTC

### ➤ Translated protein sequence:

gattgggccccgacgtcgcgatgctcccgccgcatggcgccgcccgggaattcgatttatgga  
 L G P T S H A P G R H G G R G N S I Y G  
tccatcgagatcgatgcactcttcgaaaacattgacttccaggcgaccattacccgtgct  
 S I E I D A L F E N I D F Q A T I T R A  
 cgtttcgaggaactgtgcggtgacctgttccgtggtacattgcaacccgtagagcgtgtg  
 R F E E L C G D L F R G T L Q P V E R V  
 ctgcaggatgccaagatggacaaacgcgcccgtgcacgatgtggtcctggttggtggttcc  
 L Q D A K M D K R A V H D V V L V G G S  
 acccgatatccccaaggtgatgcagcttgtgtctgactttttcggtggcaaggagcttaac  
 T R I P K V M Q L V S D F F G G K E L N  
 aagagcatcaaccccgacgagggcagtggtgacggtgctgtacaagctttcatcttg  
 K S I N P D E A V A Y G A A V Q A F I L  
 accggtggtgaagagcaagcaaacggaaggtctcctccttctggatgtcgccccgcttacg  
 T G G K S K Q T E G L L L L D V A P L T  
 ctcggcacatcgagacggcaggtggcggtgatgacggcactgatcaagcgcaacacgacgatc  
 L G I E T A G G V M T A L I K R N T T I  
 cccacaaaaagagccaaatcttctccacataactcagacaatcagcccggtgtgcacatt  
 P T K K S Q I F S T Y S D N Q P G V H I  
 caggtcttttgagggtgagcggtaccatgacgaaggactgccatctcttgggcaccttcgat  
 Q V F E G E R T M T K D C H L L G T F D  
 ctctctggaatcccccgccgaccccgcgcggtgccacagatagaagtgcaccttcgaccta



L S G I P P A P R G V P Q I E V T F D L  
gacgctaattggtattctcgagcgc

### ➤ TritrDB BLAST alignments:

> Tb927.11.11330:mRNA-p1 | transcript=Tb927.11.11330:mRNA | gene=Tb927.11.11330  
| organism=Trypanosoma brucei brucei TREU927  
| gene\_product=heat shock protein 70 | transcript\_product=heat  
shock protein 70 | location=Tb927\_11\_v5.1:3028681-3030666(+)  
| protein\_length=661 | sequence\_SO=chromosome | SO=protein\_coding  
Length=661

Score = 422 bits (1084), Expect = 6e-144, Method: Compositional matrix adjust.  
Identities = 205/205 (100%), Positives = 205/205 (100%), Gaps = 0/205 (0%)

Query	1	IEIDALFENIDFQATITRARFEELCGDLFRGTLQPVERVLQDAKMDKRAVHDVVLVGGST	60
		IEIDALFENIDFQATITRARFEELCGDLFRGTLQPVERVLQDAKMDKRAVHDVVLVGGST	
Sbjct	285	IEIDALFENIDFQATITRARFEELCGDLFRGTLQPVERVLQDAKMDKRAVHDVVLVGGST	344
Query	61	RIPKVMQLVSDFFGGKELNKSINPDEAVAYGAAVQAFILTTGGKSKQTEGLLLLDVAPLTL	120
		RIPKVMQLVSDFFGGKELNKSINPDEAVAYGAAVQAFILTTGGKSKQTEGLLLLDVAPLTL	
Sbjct	345	RIPKVMQLVSDFFGGKELNKSINPDEAVAYGAAVQAFILTTGGKSKQTEGLLLLDVAPLTL	404
Query	121	GIETAGGVMTALIKRNTTIPTKKSQIFSTYSNQPVGHIQVFEGERTMTKDCHLLGTFDL	180
		GIETAGGVMTALIKRNTTIPTKKSQIFSTYSNQPVGHIQVFEGERTMTKDCHLLGTFDL	
Sbjct	405	GIETAGGVMTALIKRNTTIPTKKSQIFSTYSNQPVGHIQVFEGERTMTKDCHLLGTFDL	464
Query	181	SGIPPAPRGVPQIEVTFDLDANGIL	205
		SGIPPAPRGVPQIEVTFDLDANGIL	
Sbjct	465	SGIPPAPRGVPQIEVTFDLDANGIL	489

## 9.14 *Tb*BBP590 sequence analysis (*Tb*927.11.11660)

GAATTCGATTTATGGATCCAAATTGTCGAGTCTGCGATCCGTGTGTGAGGAATGTAACCTC  
TGTTATTGAGCGTGCCCGTAGGAGAGAAGAGCGGTCTCACTCGCGGGAGTCTTCGCCAGC  
GTCACATCGTGTAAGGATGGCGCAGAGGCTTTACTTGTGGAGCGTTGCAAAGAGCTGGT  
GTCGGAGCTGCTCGTGTTGCGTACCGATCGACGGACTAGCACGGCTGCAATAGACCAGAG  
GCTTATGGAGAGTGAGTTGATGGTGACATCCGTACGAAATGCCATAGAGAGCCTACGAAC  
TTCCCTTGATATCGATGTACCTATCGGCACACCTGCCCTGGATTTAGTAAGGACTTCTGG  
TAGTTGTGGAGAAGCAGTTGAAGGTGCTATCAGTGAGAAAAGGCAAGTTCCGCCACCGG  
ACCGAGTTTAGCGGAACCTGGACACCTCTATCCGTACCCAGCTCGAACTCTGCGTGCTTGC  
AGTGATCAGACAATTACGTGAAAAGGACCGGGTAATTTCCGCTGCGGCAACTTTGTTGGG  
AGATGCGCTGCGTTCATGCTCCAGTGCCATTAAAGACGAAGTTACTACTGATTCGCTTTT  
AGCGGTTGCAAGCGGTACTGCTCACAGTCTGCACTCCCTATTCTCAGAGTTACAGAGTG  
TCGTGCGTTCGCTTAAACAATCGCAGCACACCGGAAGAGCTGCGTACGCGGTTGGAGGG  
GGGAGAAACAACGCAGCTGGCAACTCTGGAGGCGCAGACGGAGTCAGTTGCCACAGCGGA  
ACGGCTACTTAATGAGGCACTACAACGACAAGTTGCGGCATCAGCGCAAACAGCAAGACT  
ACAGCGGACGCTCATGGACTTCGAGGAAGCATAGCAGCTTCAACGCTTTATGCTGACGC  
TCGAGCGCAATCACTANTGAATTC

## ➤ Translated protein sequence:

**ggatccc**aaattgtcagagtctgcgatccgtgtgtgaggaatgtaactctgttattgagcgt  
 G S K L S S L R S V C E E C N S V I E R  
 gcccgtaggagagaagagcggtctcactcgcgggagtccttcgccagcgctcacatcggtga  
 A R R R E E R S H S R E S S P A S H R V  
 gaggatggcgcagaggctttacttgtggagcgttgcaaagagctggtgtcggagctgctc  
 E D G A E A L L V E R C K E L V S E L L  
 gtgttgcgtagccgatcgacggactagcacggctgcaatagaccagaggcttatggagagt  
 V L R T D R R T S T A A I D Q R L M E S  
 gagttgatggtgacatccgtacgaaatgccatagagagcctacgaacttcccttgatatac  
 E L M V T S V R N A I E S L R T S L D I  
 gatgtacctatcggcacacctgccctggatttagtaaggacttctggtagttgtggagaa  
 D V P I G T P A L D L V R T S G S C G E  
 gcagttgaaggtgctatcagtgagaaaaggcaagtttccgccaccggaccgagtttagcg  
 A V E G A I S E K R Q V S A T G P S L A  
 gaactggacacctctatccgtacccagctcgaactctgcgtgcttgacagtgatcagacaa  
 E L D T S I R T Q L E L C V L A V I R Q  
 ttacgtgaaaaggaccgggtaatttccgctgcggcaactttgttgggagatgcgctgcgt  
 L R E K D R V I S A A A T L L G D A L R  
 tcatgctccagtgccattaaagacgaagttactactgattcgcttttagcggttgcaagc  
 S C S S A I K D E V T T D S L L A V A S  
 ggtactgctcacagtctgcactccctattctcagagttacagagtgttcgtgcgtcgctt  
 G T A H S L H S L F S E L Q S V R A S L  
 aaacaatcgcagcaccacacggaagagctgcgtacgcggttgaggggggagaaacaacg  
 K Q S Q H H T E E L R T R L E G G E T T  
 cagctggcaactctggaggcgcagacggagtcagttgccacagcgggaacggctacttaac  
 Q L A T L E A Q T E S V A T A E R L L N  
 gaggcactacaacgacaagttgcggcatcagcgcgaaacagcaagactacagcggacgctc  
 E A L Q R Q V A A S A Q T A R L Q R T L  
 atggacttgcgaggaagcatagcagcttcaacgctttatgctgacg**ctcgagc**gcaatca  
 M D F A G S I A A S T L Y A D A R A Q S

## ➤ TritrDB BLAST alignments:

> [Tb927.11.10660:mRNA-p1](#) | transcript=Tb927.11.10660:mRNA | gene=Tb927.11.10660  
 | organism=Trypanosoma brucei brucei TREU927  
 | gene\_product=Basal body protein | transcript\_product=Basal  
 body protein | location=Tb927\_11\_v5.1:2840342-2856547(+)  
 | protein\_length=5401 | sequence\_SO=chromosome | SO=protein\_coding  
 Length=5401

Score = 566 bits (1458), Expect = 0.0, Method: Compositional matrix adjust.  
 Identities = 291/293 (99%), Positives = 293/293 (100%), Gaps = 0/293 (0%)

Query	1	KLSSLRSVCEECSNVIERARRREERSHSRESSPASHRVEDGAEALLVERCKELVSELLVL	60
Sbjct	3621	KLSSLRSVCEECSNVIERARRREERSHSRESSPASHRVEDGAEALLVERCKELVSELLVL	3680
Query	61	RTDRRTSTAAIDQRLMESELMVTSVRNAIESLRTSLDIDVPIGTPALDLVRTSGSCGEAV	120
Sbjct	3681	RTDRRTSTAAIDQRLMESELMVTSVRNAIE+LRTSLDIDVPIGTPALDLVRTSGSCGEAV	3740
Query	121	EGAISEKRQVSATGPSLAELDTSIRTQLELCVLAVIRQLREKDRVISAAATLLGDALRSC	180
Sbjct	3741	EGAISEKRQVSATGPSLAELDTSIRTQLELCVLAVIRQLREKDRVISAAATLLGDALRSC	3800
Query	181	SSAIKDEVTTDSLLAVASGTAHSLHSLFSELQSVRASLKQSQHHTTELRLRLEGGETTQL	240
Sbjct	3801	SSAIKDEVTTDSLLAVASGTAHSLHSLFSELQSVRASLKQSQHHTTELRLRLEGGETTQL	3860

Query	241	ATLEAQTESVATAERLLNEALQRQVAASAQTARLQRTLMDFAGSIAASTLYAD	293
		ATLEAQTESVATAERLLNEALQRQVAASAQTARLQRTLMDFAGSIAASTLYAD	
Sbjct	3861	ATLEAQTESVATAERLLNEALQRQVAASAQTARLQRTLMDFAGSIAASTLYAD	3913

## 9.15 *Tb*Protein Kinase sequence analysis (*Tb*927.9.6560)

GAATTCGATTAT GGATCCTTTCTGCTTCTCGCAGACTGAGCGGCCCGAGGGATGGGAACCG  
 CAGGACGTGAAGGTAATTGGCCGCCAGTTCCAGCCAAAGTTCTCCCGGAGAGGCTGCGA  
 AATACTGGCCAGTCCGACGACCCGGCTGCTGCAGCGCGTGACAACAGGGACGAAGCGGCT  
 GCCAGCTCCGTCAAGAGCTGCACAGCGGCACAGGAGTCTGGAGATAATGATCAAATGGTT  
 CTCAAGGCAGTAACTGCACTGAGTTCAGATACCGCTTCCACAGATCCCGAGGTGCTAGCC  
 TACCGTGAAAAGCTCATTTCGTGAACAGGAGGAGGCATGGCAAGTGGCAGAAAGCGCACACAAA  
 AAGCGGGATAAGGAAGACAAGGATGGGACGACCTCTGGACCGGCAGATGTTGATTTCATTG  
 TTTGGTCCAACCTGAAAAGAAGGAAGCTGTAAAGGAAAACAAGGCATCTGCTATTGACGAC  
 TTATTTGGTGGCATGGAGACACAACCATCACAGCCTCAAAAACCAACAACGGATGACCTC  
 TTTCTCGAGCCAATCACTAGTGAATTC

### ➤ Translated protein sequence:

gaattcgattat ggatccttctgcttctcgcagactgagcggcccgagggatgggaaccg  
 E F D Y G S F C F S Q T E R P E G W E P  
 caggacgtgaaggtaattggccgccagttccagccaaagttctcccgagagggctgcca  
 Q D V K V I G R P V P A K V L P E R L R  
 aatactggccagtcgacgacccggctgctgcagcgcgtgacaacagggacgaagcggct  
 N T G Q S D D P A A A A R D N R D E A A  
 gccagctccgtcaagagctgcacagcggcacaggagtctggagataatgatcaaattggtt  
 A S S V K S C T A A Q E S G D N D Q M V  
 ctcaaggcagtaactgcactgagttcagataccgcttccacagatcccgaggtgctagcc  
 L K A V T A L S S D T A S T D P E V L A  
 taccgtgaaaagctcatttcgtgaacaggaggaggcatggcaagtggcagaaagcgcacac  
 Y R E K L I R E Q E E A W Q V A E S A H  
 aaaaagcgggataaggaagacaaggatgggacgacctctggaccggcagatgttgattca  
 K K R D K E D K D G T T S G P A D V D S  
 ttgtttggtccaactgaaaagaaggaagctgtaaaggaaaacaaggcatctgctattgac  
 L F G P T E K K E A V K E N K A S A I D  
 gacttatttgggtggcatggagacacaaccatcacagcctcaaaaaccaacaacggatgac  
 D L F G G M E T Q P S Q P Q K P T T D D  
 ctctttt ctcgagccaatcactagtgaattcgcgccgcctgcaggctcgaccatatgggag  
 L F L E P I T S E F A A A C R S T I W E

### ➤ TritpDB BLAST alignments:

```
> Tb927.9.6560:mRNA-p1 | transcript=Tb927.9.6560:mRNA | gene=Tb927.9.6560
| organism=Trypanosoma brucei_brucei_TREU927 | gene_product=NAK
family pseudokinase, putative | transcript_product=NAK
family pseudokinase, putative | location=Tb927_09_v5.1:1077469-1079649(+)
| protein_length=726 | sequence_SO=chromosome
| SO=protein_coding
Length=726
```

Score = 348 bits (893), Expect = 9e-114, Method: Compositional matrix adjust.  
 Identities = 173/178 (97%), Positives = 174/178 (98%), Gaps = 0/178 (0%)

Query	22	GSFCFSQTERPEGWEPQDVKVI	GRPVPAKVLPERLRNTGQSDDP	AAAAARDNRDEAAASSV	81
Sbjct	339	GPFCSQTERPEGWEPQDVKVI	GRPVPAKVLPERLRNTGQSDDP	AAAAARDNRDEAAA+SV	398
Query	82	KSCTAAQESGDNDQMVLKAVT	ALSSDTASTDPEVLAYREKL	IREQEEAWQVAESAHHKRD	141
Sbjct	399	KSRTAAQESGDNDQMVLKAVT	ALSSDTASTDPEVLAYREKL	IREQEEAWQVAESAHHKRD	458
Query	142	KEDKDGTTSGPADVDSLFGP	TEKKEAVKENKASAIDDLFG	GMETQPSQPQKPTTDDL	199
Sbjct	459	KEDKDGTTSGPADVDSLFGP	TEKKEAVKENKASAIDDLFG	GMETQPSQSQKPTTDDL	516

## 9.16 *TbP25α* sequence analysis (*Tb927.4.2740*)

GAATTCACTAGTGATTGCG**GGATCCTGGAAGCTGTCTTCTACGCATTTGCATCCTTCGGCAC**  
**GGCGCCACGAAGGAGATGGACAGTGCCCATTTCTCCAAAATGTTAAAGGAGGCTAAAATCA**  
**TTGGCAAAACCTTCACCTCTACAGATGCCGACCTCCTCTTCAACAAGATCAAGGCGAAGGGC**  
**GCCCCGTAAAATCACCTTCACGGAGTTTAAACACAAGAGCCCTCCCTGATATTGCCACCAAGTT**  
**GAAGATGACACCCGAGCAGGTGGCTGAAATTCTCACGAAGGCATCACCCGCCTCCAATTCCA**  
**CAAAAGCAGAAGCTGTTAAGTTCCATGACGACAAAAATCTCTACACGGGCGTCTACAAGGCG**  
**GGAGGCCCCACAAATGTGGATCGTAACGCCGGATCCCTTTCAGGGGTTGTCGACCGCCGTGT**  
**GGATCAGGTCGATGTGCTCGAGCGCAATCGAATTC**

### ➤ Translated protein sequence:

ttgggagctctcccatatgggtcgacctgcaggcggccgcgaattcactagtgattgcg**gg**  
L G A L P Y G R P A G G R E F T S D C G  
**atcc**tggaagctgtcttctacgcatttgcacatccttcggcacggcgccacgaaggagatg  
I L E A V F Y A F A S F G T A P T K E M  
gacagtgcccatTTCTCCAAAATGTTAAAGGAGGCTAAAATCATTGGCAAAACCTTCACC  
D S A H F S K M L K E A K I I G K T F T  
tctacagatgccgacctcctcttcaacaagatcaaggcgaagggcgcccgtaaaatcacc  
S T D A D L L F N K I K A K G A R K I T  
ttcacggagtttaacacaagagccctccctgatattgccaccaagttgaagatgacaccc  
F T E F N T R A L P D I A T K L K M T P  
gagcaggtggctgaaattctcacgaaggcatcacccgcctccaattccacaaaagcagaa  
E Q V A E I L T K A S P A S N S T K A E  
gctgttaagttccatgacgacaaaaatctctacacgggCGTCTACAAGGCGGGAGGCCCC  
A V K F H D D K N L Y T G V Y K A G G P  
acaaatgtggatcgtaacgccggatccctttcaggggttgctcgaccgccgtgtggatcag  
T N V D R N A G S L S G V V D R R V D Q  
gtcgatgtg**ctcgag**cgcaatcgaaattcccgcgcccgccatggcggccgggagcatgcga  
V D V L E R N R I P A A A M A A G S M R

### ➤ TritrpDB BLAST alignments:

> [Tb927.4.2740:mRNA-p1](#) | transcript=Tb927.4.2740:mRNA | gene=Tb927.4.2740

| organism=Trypanosoma\_brucei\_brucei\_TREU927 | gene\_product=p25-alpha,  
putative | transcript\_product=p25-alpha,  
putative | location=Tb927\_04\_v5.1:723054-723506(+) | protein\_length=150  
| sequence\_SO=chromosome | SO=protein\_coding  
Length=150

Score = 286 bits (732), Expect = 1e-99, Method: Compositional matrix  
adjust.

Identities = 140/141 (99%), Positives = 141/141 (100%), Gaps = 0/141 (0%)

```
Query 1 EAVFYAFASFGTAPTKEMDSAHSKMLKEAKIIGKTFTSTDADLLFNKIKAKGARKITFT 60
      EAVFYAFASFGTAPTKEMD+AHFSKMLKEAKIIGKTFTSTDADLLFNKIKAKGARKITFT
Sbjct 2 EAVFYAFASFGTAPTKEMDNAHSKMLKEAKIIGKTFTSTDADLLFNKIKAKGARKITFT 61

Query 61 EFNTRALPDIATKLKMTPEQVAEILTKASPASNSTKAEAVKFHDDKNLYTGVYKAGGPTN 120
      EFNTRALPDIATKLKMTPEQVAEILTKASPASNSTKAEAVKFHDDKNLYTGVYKAGGPTN
Sbjct 62 EFNTRALPDIATKLKMTPEQVAEILTKASPASNSTKAEAVKFHDDKNLYTGVYKAGGPTN 121

Query 121 VDRNAGSLSGVVD RRV DQVDV 141
      VDRNAGSLSGVVD RRV DQVDV
Sbjct 122 VDRNAGSLSGVVD RRV DQVDV 142
```

## 9.17 *Tb*BARTL1 sequence analysis (*Tb*927.10.5810)

GAATTCAGTAGTGATTGCGGGATCCTTTTACAGTTCACGCATTTCGCCAATGTGGCGTACCCCC  
GTTGACAACCTTCATTGACGAGAGGTGTGCCCTCTCCACGGACGATGAAGAGATGAAGGTCGA  
GCAGACAGAGGTACACATGCAATTTTCGAAAAATTACCGATGACATTCTTTCTAAATACGTGA  
AGGAGCTAGGCATAACGCTAGAGGACGCGCTGAATGCTGTGGTGAAGTCCATGGACACCGCG  
AGCCAGACGAACAGGCTTGGCAAGAAGTTCATGCAGGAAATCTTCTATATCGAGGATTTCCC  
CACCTTCCACAAGATGATGGTGCGACGCAACATCGAGCTCGACATCCTTGCTCAGTGTGAAA  
TCAGTCAAAAACGTGAGAACAGTGGTGCAAAAACGTGCAAAATGATGAGGAGGAGGCGATGAGG  
CTCGCTATCGAGGCTTCTCTTAACGATGAGGAGAAGACCCGCCGTCTAATGGAGTTGGAGGA  
TCTGCAGCTACAGGCTCGAGCGCAATCGAATTC

### ➤ Translated protein sequence:

tgcaggcgccgcgaattcactagtgattgcgggatcctttacagttcacgcatttcgcca  
C R R P R I H - - L R D P L Q F T H S P  
atgtggcgtagcccggttgacaacttcattgacgagaggtgtgccctctccacggacgat  
M W R T P V D N F I D E R C A L S T D D  
gaagagatgaaggtcgagcagacagaggtacacatgcaatttcgaaaaattaccgatgac  
E E M K V E Q T E V H M Q F R K I T D D  
attctttctaaatacgtgaaggagctaggcataacgctagaggacgcgctgaatgctgtg  
I L S K Y V K E L G I T L E D A L N A V  
gtgaactccatggacaccgagccagacgaacaggccttgcaagaagttcatgcaggaa  
V N S M D T A S Q T N R L G K K F M Q E  
atcttctatatcgaggatttccccaccttcacaagatgatggtgacgacgaacatcgag  
I F Y I E D F P T F H K M M V R R N I E  
ctcgacatccttgctcagtgatgaaatcagtcaaaaacgtgagaacagtggtgcaaaaact  
L D I L A Q C E I S Q K R E N S G A K T  
gcaaatgatgaggaggaggcgatgaggctcgctatcgaggcttctcttaacgatgaggag  
A N D E E E A M R L A I E A S L N D E E  
aagacccgccgtctaattggaggttgaggatctgcagctacaggctcgagcgcaatcgaat

K T R R L M E L E D L Q L Q A R A Q S N

### ➤ TritrpDB BLAST alignments:

```
> Tb927.10.5810:mRNA-p1 | transcript=Tb927.10.5810:mRNA |
gene=Tb927.10.5810
| organism=Trypanosoma brucei brucei TREU927 |
gene_product=The ARF-like 2 binding protein BART, putative
| transcript_product=The ARF-like 2 binding protein BART, putative
| location=Tb927_10_v5.1:1441554-1442855(-) | protein_length=433
| sequence_SO=chromosome | SO=protein_coding
Length=433

Score = 338 bits (867), Expect = 9e-116, Method: Compositional
matrix adjust.
Identities = 160/161 (99%), Positives = 160/161 (99%), Gaps = 0/161
(0%)

Query 1 LQFTHSPMWRTPVDNFIDERCALSTDDEEMKVEQTEVHMQFRKITDDILSKYVKELGITL 60
Sbjct 11 LQFTHSPMWRTPVDNFIDERCALSTDDEEMKVEQTEVHMQFRKITDDILSKYVKELGITL 70

Query 61 EDALNAVVNNSMDTASQTNRLGKKFMQEIFYIEDFPTFHKMMVRRNIELDILAQCEISQKR 120
Sbjct 71 EDALNAVVNNSMDTASQTNRLGKKFMQEIFYIEDFPTFHKMMVRRNIELDILAQCEISQKR 130

Query 121 ENSGAKTANDEEEAMRLAIEASLNDEEKTRRLMELEDLQLQ 161
Sbjct 131 ENSGAKTANDEEEAMRLAIEASLNDEEKTRRLMELEDLQLQ 171
```

## 9.18 *Tb*TCP-1-ε sequence analysis (*Tb*927.11.14250)

GAATTCAGTGTGATTGCGGGATCCTCTCGCATCTCTGAGGGGTTTGAGAAGGCCTGTGA  
AATTGCCTGTAAGCGGTTGGAGGAGATCGCAGACACCGTGCCCGTCAGCCGTGAGGAATA  
CAGTTATTTGCTGCAAACCGCGCGCTCAACACTCAACTCGAAGGTGGTAAACCGTGACCG  
TGATCGGTTGGCAAAGATATGTGTTGATGCGGTTCTTTTCGGTCGCAGACATGGAACGCCG  
TGACGTCAATCTCGATCTAATTAAGATGGAGGGTAAAGTCGGTGGATGCCTTGAGGAAAC  
ATGCTTAGTGAAACGGCATCGTCATTGACAAGGACTTCTCCCATCCTCAAATGCCCAAGGT  
GCTGAAGAATCCAAAGATTGCCATCCTCACATGCCCATTTGAGCCCCCGAAACCCAAAAC  
TAAACACACCGTGACATTTGAGTGCCGAGCACATGAAAGAAATTCACGAGCAAGAGCA  
GGAGTACTTCCGCACTCGAGCGCAATCGAATTC

### ➤ Translated protein sequence:

gtcgacctgcaggcgccgcgaattcactagtgtgattgcgggatccttctcgcatctctgag  
V D L Q A A A N S L V I A G S S R I S E  
gggtttgagaaggcctgtgaaattgcctgtaagcgggttgaggagatcgacacacggtg  
G F E K A C E I A C K R L E E I A D T V  
cccgtcagccgtgaggaatacagttatgtgctgcaaaccgcgcgctcaactcaactcg  
P V S R E E Y S Y L L Q T A R S T L N S  
aaggtggtaaaccgtgacggtgatcggttgcaaagatatgtgttgatgcggttctttcg  
K V V N R D R D R L A K I C V D A V L S  
gtcgacagatggaacgccgtgacgtcaatctcgatctaattaagatggagggtaaagtc

V A D M E R R D V N L D L I K M E G K V  
 ggtggatgccttgaggaaacatgcttagtgaacggcatcgatcattgacaaggacttctcc  
 G G C L E E T C L V N G I V I D K D F S  
 catcctcaaagtcccaagggtgctgaagaatccaaagattgccatcctcacatgccattt  
 H P Q M P K V L K N P K I A I L T C P F  
 gagcccccgaaacccaaaactaaacacaccgtgcacatttcgagtgccgagcacatgaaa  
 E P P K P K T K H T V H I S S A E H M K  
 gaaattcacgagcaagagcaggagtacttccgca~~ctcgag~~cgcaatcgaattcccgcggc  
 E I H E Q E Q E Y F R T R A Q S N S R G

### ➤ TritrDB BLAST alignments:

```
> Tb927.11.14250:mRNA-p1 | transcript=Tb927.11.14250:mRNA |
gene=Tb927.11.14250
| organism=Trypanosoma brucei brucei TREU927
| gene_product=T-complex protein 1, epsilon subunit, putative
| transcript_product=T-complex protein 1, epsilon subunit,
putative | location=Tb927_11_v5.1:3793465-3795081(+) |
protein_length=538
| sequence_SO=chromosome | SO=protein_coding
Length=538

Score = 327 bits (839), Expect = 1e-110, Method: Compositional
matrix adjust.
Identities = 156/156 (100%), Positives = 156/156 (100%), Gaps =
0/156 (0%)

Query 1 SRISEGFEEKACEIACKRLEEIADTVPVSRREEYSYLLQTARSTLNSKVVNRDRDLAKICV 60
Sbjct 127 SRISEGFEEKACEIACKRLEEIADTVPVSRREEYSYLLQTARSTLNSKVVNRDRDLAKICV 186

Query 61 DAVLSVADMERRDVNLDLIKMEGKVGCGLEETCLVNGIVIDKDFSHQMPKVLKNPKIAI 120
Sbjct 187 DAVLSVADMERRDVNLDLIKMEGKVGCGLEETCLVNGIVIDKDFSHQMPKVLKNPKIAI 246

Query 121 LTCPFEPKPKTKHTVHISAEHMKEIHEQEYFR 156
Sbjct 247 LTCPFEPKPKTKHTVHISAEHMKEIHEQEYFR 282
```

## 9.19 *Tb*VHS sequence analysis (*Tb*927.11.10770)

GAATTCAGTAGTGATTGCG~~GGATCCT~~ATTGTGCTGGAGTCGCTTGTCAAAAATTGTAACACG  
 AAGTTGCACACGGAAGTCGCTGCGCAGAAAGGCATTGTGAAGGAAGTGTACAACATTGCCAC  
 CCGCAGCGCTACCAGTGAGAAGGAGTGCTTAGCGAAGGAAGCAGCTTTGGCGCTCATTCTTA  
 ATTTCTCTGTTTGGTTTGCAGGGCATCCCAACAGCCGGTTAAAGTTCTTGACGTCAGTCGCT  
 GAGGCTGTTTCGACGGGCTGTAGGGCCCAATGCATTTGATGGAATCCAGCCCGACATTGATAC  
 GCGACTCTCCATGGCAGTAGGCCCAACGGGCCAGCGCACCCGTCCCCGGCGAAGCAACCGA  
 ACAAAATCAATACACCACCCACCAGAACTGGACTCCCGCCGGGAGCACACGTGGTAGACGCA  
 ATTGGCATCGTTTTACCAACGGACGAGGA~~CTCGAG~~CGCAATCGAATTC

➤ **Translated protein sequence:**

cg**ggatccc**tattgtgctggagtcgcttgtaacacgaagttgcacacggaa  
R D P I V L E S L V K N C N T K L H T E  
gtcgtgctgcgcagaaaggcattgtgaaggaaactgtacaacattgccacccgcagcgctacc  
V A A Q K G I V K E L Y N I A T R S A T  
agtgagaaggagtgcttagcgaaggaagcagctttggcgctcattcttaatttctctgtt  
S E K E C L A K E A A L A L I L N F S V  
tggtttgcagggcatcccaacagccggttaaagttcttgacgtcagtcgctgaggctgtt  
W F A G H P N S R L K F L T S V A E A V  
cgacgggctgtagggcccaatgcatttgatggaatccagcccgcacattgatacgcgactc  
R R A V G P N A F D G I Q P D I D T R L  
tccatggcagtagggcccaacggggccagcgcacccgtcccccggaagcaaccgaacaaa  
S M A V G P T G Q R T R P P A K Q P N K  
tcaatacaccacccaccagaaactggactcccgcggggagcacacgtggtagacgcaatt  
S I H H P P E T G L P P G A H V V D A I  
ggcatcgttttaccaacggacgagga**ctcgag**cgcaatcgaattccccgcggccgcatgg  
G I V L P T D E D S S A I E F P R P P W

➤ **TritrpDB BLAST alignments:**

> Tb927.11.10770:mRNA-p1 | transcript=Tb927.11.10770:mRNA | gene=Tb927.11.10770  
| organism=Trypanosoma\_brucei\_brucei\_TREU927  
| gene\_product=VHS domain containing protein, putative | transcript\_product=VHS  
domain containing protein, putative |  
location=Tb927\_11\_v5.1:2883543-2884919(+) | protein\_length=458  
| sequence\_SO=chromosome | SO=protein\_coding  
Length=458

Score = 303 bits (777), Expect = 1e-102, Method: Compositional matrix adjust.  
Identities = 145/145 (100%), Positives = 145/145 (100%), Gaps = 0/145 (0%)

Query	1	IVLESLVKNCNTKLHTEVAAQKGIVKELYNIATRSATSEKECLAKEAALALILNFSVWFA	60
		IVLESLVKNCNTKLHTEVAAQKGIVKELYNIATRSATSEKECLAKEAALALILNFSVWFA	
Sbjct	80	IVLESLVKNCNTKLHTEVAAQKGIVKELYNIATRSATSEKECLAKEAALALILNFSVWFA	139
Query	61	GHPNSRLKFLTSVAEAVRRVAVGNPAFDGIQPDIDTRLSMAVGPTGQRTTRPPAKQPKNKSIH	120
		GHPNSRLKFLTSVAEAVRRVAVGNPAFDGIQPDIDTRLSMAVGPTGQRTTRPPAKQPKNKSIH	
Sbjct	140	GHPNSRLKFLTSVAEAVRRVAVGNPAFDGIQPDIDTRLSMAVGPTGQRTTRPPAKQPKNKSIH	199
Query	121	HPPETGLPPGAHVVD AIGIVLPTDE	145
		HPPETGLPPGAHVVD AIGIVLPTDE	
Sbjct	200	HPPETGLPPGAHVVD AIGIVLPTDE	224



## References

- Absalon, S., T. Blisnick, M. Bonhivers, L. Kohl, N. Cayet, G. Toutirais, J. Buisson, D. Robinson, and P. Bastin, 2008a, Flagellum elongation is required for correct structure, orientation and function of the flagellar pocket in *Trypanosoma brucei*: *J Cell Sci*, v. 121, p. 3704-16.
- Absalon, S., T. Blisnick, L. Kohl, G. Toutirais, G. Doré, D. Julkowska, A. Tavenet, and P. Bastin, 2008b, Intraflagellar transport and functional analysis of genes required for flagellum formation in trypanosomes.: *Mol Biol Cell*, v. 19, p. 929-44.
- Adams, M., U. M. Smith, C. V. Logan, and C. A. Johnson, 2008, Recent advances in the molecular pathology, cell biology and genetics of ciliopathies.: *J Med Genet*, v. 45, p. 257-67.
- Adam, R. D., 2001, Biology of *Giardia lamblia*: *Clin Microbiol Rev*, v. 14, p. 447-75.
- Adhiambo, C., J. D. Forney, D. J. Asai, and J. H. LeBowitz, 2005, The two cytoplasmic dynein-2 isoforms in *Leishmania mexicana* perform separate functions: *Mol Biochem Parasitol*, v. 143, p. 216-25.
- Allen, C. L., D. Goulding, and M. C. Field, 2003, Clathrin-mediated endocytosis is essential in *Trypanosoma brucei*: *EMBO J*, v. 22, p. 4991-5002.
- Alsford, S., D. J. Turner, S. O. Obado, A. Sanchez-Flores, L. Glover, M. Berriman, C. Hertz-Fowler, and D. Horn, 2011, High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome: *Genome Res*, v. 21, p. 915-24.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, 1990, Basic local alignment search tool: *J Mol Biol*, v. 215, p. 403-10.
- Amor, J. C., J. R. Horton, X. Zhu, Y. Wang, C. Sullards, D. Ringe, X. Cheng, and R. A. Kahn, 2001, Structures of yeast ARF2 and ARL1: distinct roles for the N terminus in the structure and function of ARF family GTPases: *J Biol Chem*, v. 276, p. 42477-84.
- Anderson, R. G., 1972, The three-dimensional structure of the basal body from the rhesus monkey oviduct: *J Cell Biol*, v. 54, p. 246-65.
- Andre, J., L. Kerry, X. Qi, E. Hawkins, K. Drizyte, M. L. Ginger, and P. G. McKean, 2013, An alternative model for the role of RP2 in flagellum assembly in the African trypanosome: *J Biol Chem*.
- Andre, J., L. Kerry, X. Qi, E. Hawkins, K. Drizyte, M. L. Ginger, and P. G. McKean, 2014, An alternative model for the role of RP2 protein in flagellum assembly in the African trypanosome: *J Biol Chem*, v. 289, p. 464-75.
- André, J., S. Harrison, K. Towers, X. Qi, S. Vaughan, P. G. McKean, and M. L. Ginger, 2013, Tubulin-binding cofactor C domain-containing protein TBCCD1 orchestrates cytoskeletal filament formation: *J Cell Sci*.
- Andréasson, C., J. Fiaux, H. Rampelt, M. P. Mayer, and B. Bukau, 2008, Hsp110 is a nucleotide-activated exchange factor for Hsp70: *J Biol Chem*, v. 283, p. 8877-84.
- Arts, H. H., D. Doherty, S. E. van Beersum, M. A. Parisi, S. J. Letteboer, N. T. Gorden, T. A. Peters, T. Märker, K. Voeselek, A. Kartono, H. Ozyurek, F. M. Farin, H. Y. Kroes, U. Wolfrum, H. G. Brunner, F. P. Cremers, I. A. Glass, N. V. Knoers, and R. Roepman, 2007, Mutations in the gene encoding the basal body protein RPGRIP1L, a nephrocystin-4 interactor, cause Joubert syndrome: *Nat Genet*, v. 39, p. 882-8.
- Aslett, M., C. Aurrecochea, M. Berriman, J. Brestelli, B. P. Brunk, M. Carrington, D. P. Depledge, S. Fischer, B. Gajria, X. Gao, M. J. Gardner, A. Gingle, G. Grant, O. S. Harb, M. Heiges, C. Hertz-Fowler, R. Houston, F. Innamorato, J. Iodice, J. C. Kissinger, E. Kraemer, W. Li, F. J. Logan, J. A. Miller, S. Mitra, P. J. Myler, V. Nayak, C. Pennington, I. Phan, D. F. Pinney, G. Ramasamy, M. B. Rogers, D. S. Roos, C. Ross, D. Sivam, D. F.

- Smith, G. Srinivasamoorthy, C. J. Stoeckert, S. Subramanian, R. Thibodeau, A. Tivey, C. Treatman, G. Velarde, and H. Wang, 2010, TriTrypDB: a functional genomic resource for the Trypanosomatidae: *Nucleic Acids Res*, v. 38, p. D457-62.
- Assimon, V. A., A. T. Gillies, J. N. Rauch, and J. E. Gestwicki, 2013, Hsp70 protein complexes as drug targets: *Curr Pharm Des*, v. 19, p. 404-17.
- Avidor-Reiss, T., A. M. Maer, E. Koundakjian, A. Polyanovsky, T. Keil, S. Subramaniam, and C. S. Zuker, 2004, Decoding cilia function: defining specialized genes required for compartmentalized cilia biogenesis: *Cell*, v. 117, p. 527-39.
- Awata, J., S. Takada, C. Standley, K. F. Lehtreck, K. D. Bellvé, G. J. Pazour, K. E. Fogarty, and G. B. Witman, 2014, NPHP4 controls ciliary trafficking of membrane proteins and large soluble proteins at the transition zone: *J Cell Sci*, v. 127, p. 4714-27.
- Azem, A., W. Oppliger, A. Lustig, P. Jenö, B. Feifel, G. Schatz, and M. Horst, 1997, The mitochondrial hsp70 chaperone system. Effect of adenine nucleotides, peptide substrate, and mGrpE on the oligomeric state of mhsp70: *J Biol Chem*, v. 272, p. 20901-6.
- Badano, J. L., N. Mitsuma, P. L. Beales, and N. Katsanis, 2006, The ciliopathies: an emerging class of human genetic disorders: *Annu Rev Genomics Hum Genet*, v. 7, p. 125-48.
- Baiady, N., P. Padala, B. Mashahreh, E. Cohen-Kfir, E. A. Todd, K. E. Du Pont, C. E. Berndsen, and R. Wiener, 2016, The Vps27/Hrs/STAM (VHS) Domain of the Signal-transducing Adaptor Molecule (STAM) Directs Associated Molecule with the SH3 Domain of STAM (AMSH) Specificity to Longer Ubiquitin Chains and Dictates the Position of Cleavage: *J Biol Chem*, v. 291, p. 2033-42.
- Bangs, J. D., L. Uyetake, M. J. Brickman, A. E. Balber, and J. C. Boothroyd, 1993, Molecular cloning and cellular localization of a BiP homologue in *Trypanosoma brucei*. Divergent ER retention signals in a lower eukaryote: *J Cell Sci*, v. 105 ( Pt 4), p. 1101-13.
- Bao, X. Q., and G. T. Liu, 2009, Induction of overexpression of the 27- and 70-kDa heat shock proteins by bicyclol attenuates concanavalin A-Induced liver injury through suppression of nuclear factor-kappaB in mice: *Mol Pharmacol*, v. 75, p. 1180-8.
- Barbelanne, M., D. Hossain, D. P. Chan, J. Peränen, and W. Y. Tsang, 2015, Nephrocystin proteins NPHP5 and Cep290 regulate BBSome integrity, ciliary trafficking and cargo delivery: *Hum Mol Genet*, v. 24, p. 2185-200.
- Barker, A. R., K. S. Renzaglia, K. Fry, and H. R. Dawe, 2014, Bioinformatic analysis of ciliary transition zone proteins reveals insights into the evolution of ciliopathy networks: *BMC Genomics*, v. 15, p. 531.
- Barr, M. M., J. DeModena, D. Braun, C. Q. Nguyen, D. H. Hall, and P. W. Sternberg, 2001, The *Caenorhabditis elegans* autosomal dominant polycystic kidney disease gene homologs *lov-1* and *pkd-2* act in the same pathway: *Curr Biol*, v. 11, p. 1341-6.
- Barr, M. M., and P. W. Sternberg, 1999, A polycystic kidney-disease gene homologue required for male mating behaviour in *C. elegans*: *Nature*, v. 401, p. 386-9.
- Bartolini, F., A. Bhamidipati, S. Thomas, U. Schwahn, S. A. Lewis, and N. J. Cowan, 2002, Functional overlap between retinitis pigmentosa 2 protein and the tubulin-specific chaperone cofactor C: *J Biol Chem*, v. 277, p. 14629-34.
- Bastin, P., and K. Gull, 1999, Assembly and function of complex flagellar structures illustrated by the paraflagellar rod of trypanosomes: *Protist*, v. 150, p. 113-23.
- Bastin, P., T. Sherwin, and K. Gull, 1998, Paraflagellar rod is vital for trypanosome motility: *Nature*, v. 391, p. 548.
- Bastin, P., T. H. MacRae, S. B. Francis, K. R. Matthews, and K. Gull, 1999, Flagellar morphogenesis: protein targeting and assembly in the paraflagellar rod of trypanosomes: *Mol Cell Biol*, v. 19, p. 8191-200.

- Beghin, A., S. Honore, C. Messana, E. L. Matera, J. Aim, S. Burlinchon, D. Braguer, and C. Dumontet, 2007, ADP ribosylation factor like 2 (Arl2) protein influences microtubule dynamics in breast cancer cells: *Exp Cell Res*, v. 313, p. 473-85.
- Berbari, N. F., A. K. O'Connor, C. J. Haycraft, and B. K. Yoder, 2009, The primary cilium as a complex signaling center.: *Curr Biol*, v. 19, p. R526-35.
- Berriman, M., E. Ghedin, C. Hertz-Fowler, G. Blandin, H. Renauld, D. C. Bartholomeu, N. J. Lennard, E. Caler, N. E. Hamlin, B. Haas, U. Böhme, L. Hannick, M. A. Aslett, J. Shallom, L. Marcello, L. Hou, B. Wickstead, U. C. Alsmark, C. Arrowsmith, R. J. Atkin, A. J. Barron, F. Bringaud, K. Brooks, M. Carrington, I. Cherevach, T. J. Chillingworth, C. Churcher, L. N. Clark, C. H. Corton, A. Cronin, R. M. Davies, J. Doggett, A. Djikeng, T. Feldblyum, M. C. Field, A. Fraser, I. Goodhead, Z. Hance, D. Harper, B. R. Harris, H. Hauser, J. Hostetler, A. Ivens, K. Jagels, D. Johnson, J. Johnson, K. Jones, A. X. Kerhornou, H. Koo, N. Larke, S. Landfear, C. Larkin, V. Leech, A. Line, A. Lord, A. Macleod, P. J. Mooney, S. Moule, D. M. Martin, G. W. Morgan, K. Mungall, H. Norbertczak, D. Ormond, G. Pai, C. S. Peacock, J. Peterson, M. A. Quail, E. Rabinowitsch, M. A. Rajandream, C. Reitter, S. L. Salzberg, M. Sanders, S. Schobel, S. Sharp, M. Simmonds, A. J. Simpson, L. Tallon, C. M. Turner, A. Tait, A. R. Tivey, S. Van Aken, D. Walker, D. Wanless, S. Wang, B. White, O. White, S. Whitehead, J. Woodward, J. Wortman, M. D. Adams, T. M. Embley, K. Gull, E. Ullu, J. D. Barry, A. H. Fairlamb, F. Opperdoes, B. G. Barrell, J. E. Donelson, N. Hall, C. M. Fraser, et al., 2005, The genome of the African trypanosome *Trypanosoma brucei*.: *Science*, v. 309, p. 416-22.
- Bhagat, L., V. P. Singh, R. K. Dawra, and A. K. Saluja, 2008, Sodium arsenite induces heat shock protein 70 expression and protects against secretagogue-induced trypsinogen and NF-kappaB activation: *J Cell Physiol*, v. 215, p. 37-46.
- Bhamidipati, A., S. A. Lewis, and N. J. Cowan, 2000, ADP ribosylation factor-like protein 2 (Arl2) regulates the interaction of tubulin-folding cofactor D with native tubulin: *J Cell Biol*, v. 149, p. 1087-96.
- Bhogaraju, S., B. D. Engel, and E. Lorentzen, 2013, Intraflagellar transport complex structure and cargo interactions: *Cilia*, v. 2, p. 10.
- Bhowmick, R., M. Li, J. Sun, S. A. Baker, C. Insinna, and J. C. Besharse, 2009, Photoreceptor IFT complexes containing chaperones, guanylyl cyclase 1 and rhodopsin: *Traffic*, v. 10, p. 648-63.
- Bialas, N. J., P. N. Inglis, C. Li, J. F. Robinson, J. D. Parker, M. P. Healey, E. E. Davis, C. D. Inglis, T. Toivonen, D. C. Cottell, O. E. Blacque, L. M. Quarmby, N. Katsanis, and M. R. Leroux, 2009, Functional interactions between the ciliopathy-associated Meckel syndrome 1 (MKS1) protein and two novel MKS1-related (MKSR) proteins: *J Cell Sci*, v. 122, p. 611-24.
- Blacque, O. E., M. J. Reardon, C. Li, J. McCarthy, M. R. Mahjoub, S. J. Ansley, J. L. Badano, A. K. Mah, P. L. Beales, W. S. Davidson, R. C. Johnsen, M. Audeh, R. H. Plasterk, D. L. Baillie, N. Katsanis, L. M. Quarmby, S. R. Wicks, and M. R. Leroux, 2004, Loss of *C. elegans* BBS-7 and BBS-8 protein function results in cilia defects and compromised intraflagellar transport: *Genes Dev*, v. 18, p. 1630-42.
- Blisnick, T., J. Buisson, S. Absalon, A. Marie, N. Cayet, and P. Bastin, 2014, The intraflagellar transport dynein complex of trypanosomes is made of a heterodimer of dynein heavy chains and of light and intermediate chains of distinct functions: *Mol Biol Cell*, v. 25, p. 2620-33.
- Bloch, M. A., and K. A. Johnson, 1995, Identification of a molecular chaperone in the eukaryotic flagellum and its localization to the site of microtubule assembly: *J Cell Sci*, v. 108 ( Pt 11), p. 3541-5.
- Bloodgood, R. A., 2010, Sensory reception is an attribute of both primary cilia and motile cilia.: *J Cell Sci*, v. 123, p. 505-9.

- Boldt, K., J. van Reeuwijk, Q. Lu, K. Koutroumpas, T. M. Nguyen, Y. Texier, S. E. van Beersum, N. Horn, J. R. Willer, D. A. Mans, G. Dougherty, I. J. Lamers, K. L. Coene, H. H. Arts, M. J. Betts, T. Beyer, E. Bolat, C. J. Gloeckner, K. Haidari, L. Hetterschijt, D. Iaconis, D. Jenkins, F. Klose, B. Knapp, B. Latour, S. J. Letteboer, C. L. Marcelis, D. Mitic, M. Morleo, M. M. Oud, M. Riemersma, S. Rix, P. A. Terhal, G. Toedt, T. J. van Dam, E. de Vrieze, Y. Wissinger, K. M. Wu, G. Apic, P. L. Beales, O. E. Blacque, T. J. Gibson, M. A. Huynen, N. Katsanis, H. Kremer, H. Omran, E. van Wijk, U. Wolfrum, F. Kepes, E. E. Davis, B. Franco, R. H. Giles, M. Ueffing, R. B. Russell, R. Roepman, and U. K. R. D. Group, 2016, An organelle-specific protein landscape identifies novel diseases and molecular mechanisms: *Nat Commun*, v. 7, p. 11491.
- Bonhivers, M., S. Nowacki, N. Landrein, and D. R. Robinson, 2008, Biogenesis of the trypanosome endo-exocytotic organelle is cytoskeleton mediated: *PLoS Biol*, v. 6, p. e105.
- Boulon, S., N. Marmier-Gourrier, B. Pradet-Balade, L. Wurth, C. Verheggen, B. E. Jady, B. Rothe, C. Pescia, M. C. Robert, T. Kiss, B. Bardoni, A. Krol, C. Branlant, C. Allmang, E. Bertrand, and B. Charpentier, 2008, The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery: *J Cell Biol*, v. 180, p. 579-95.
- Bouskila, M., N. Esoof, L. Gay, E. H. Fang, M. Deak, M. J. Begley, L. C. Cantley, A. Prescott, K. G. Storey, and D. R. Alessi, 2011, TTBK2 kinase substrate specificity and the impact of spinocerebellar-ataxia-causing mutations on expression, activity, localization and development: *Biochem J*, v. 437, p. 157-67.
- Brazelton, W. J., C. D. Amundsen, C. D. Silflow, and P. A. Lefebvre, 2001, The bld1 mutation identifies the *Chlamydomonas* *osm-6* homolog as a gene required for flagellar assembly: *Curr Biol*, v. 11, p. 1591-4.
- Briggs, L. J., J. A. Davidge, B. Wickstead, M. L. Ginger, and K. Gull, 2004a, More than one way to build a flagellum: comparative genomics of parasitic protozoa: *Curr Biol*, v. 14, p. R611-2.
- Briggs, L. J., P. G. McKean, A. Baines, F. Moreira-Leite, J. Davidge, S. Vaughan, and K. Gull, 2004b, The flagella connector of *Trypanosoma brucei*: an unusual mobile transmembrane junction: *J Cell Sci*, v. 117, p. 1641-51.
- Broadhead, R., H. R. Dawe, H. Farr, S. Griffiths, S. R. Hart, N. Portman, M. K. Shaw, M. L. Ginger, S. J. Gaskell, P. G. McKean, and K. Gull, 2006, Flagellar motility is required for the viability of the bloodstream trypanosome: *Nature*, v. 440, p. 224-7.
- Brown, J. M., D. A. Cochran, B. Craige, T. Kubo, and G. B. Witman, 2015, Assembly of IFT trains at the ciliary base depends on IFT74: *Curr Biol*, v. 25, p. 1583-93.
- Brugerolle, G., and J. P. Mignot, 2003, The rhizoplast of chrysomonads, a basal body-nucleus connector that polarises the dividing spindle: *Protoplasma*, v. 222, p. 13-21.
- Brun, R., and Schonenberger, 1979, Cultivation and in vitro cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Short communication: *Acta Trop*, v. 36, p. 289-92.
- Buisson, J., N. Chenouard, T. Lagache, T. Blisnick, J. C. Olivo-Marin, and P. Bastin, 2013, Intraflagellar transport proteins cycle between the flagellum and its base: *J Cell Sci*, v. 126, p. 327-38.
- Burckle, C., H. M. Gaude, C. Vesque, F. Silbermann, R. Salomon, C. Jeanpierre, C. Antignac, S. Saunier, and S. Schneider-Maunoury, 2011, Control of the Wnt pathways by nephrocystin-4 is required for morphogenesis of the zebrafish pronephros: *Hum Mol Genet*, v. 20, p. 2611-27.
- Burger, A., M. H. Ludewig, and A. Boshoff, 2014, Investigating the Chaperone Properties of a Novel Heat Shock Protein, Hsp70.c, from *Trypanosoma brucei*: *J Parasitol Res*, v. 2014, p. 172582.

- Cantagrel, V., J. L. Silhavy, S. L. Bielas, D. Swistun, S. E. Marsh, J. Y. Bertrand, S. Audollent, T. Attié-Bitach, K. R. Holden, W. B. Dobyns, D. Traver, L. Al-Gazali, B. R. Ali, T. H. Lindner, T. Caspary, E. A. Otto, F. Hildebrandt, I. A. Glass, C. V. Logan, C. A. Johnson, C. Bennett, F. Brancati, E. M. Valente, C. G. Woods, J. G. Gleeson, and I. J. S. R. D. S. Group, 2008, Mutations in the cilia gene ARL13B lead to the classical form of Joubert syndrome: *Am J Hum Genet*, v. 83, p. 170-9.
- Cardenas-Rodriguez, M., and J. L. Badano, 2009, Ciliary biology: understanding the cellular and genetic basis of human ciliopathies.: *Am J Med Genet C Semin Med Genet*, v. 151C, p. 263-80.
- Caspary, T., C. E. Larkins, and K. V. Anderson, 2007, The graded response to Sonic Hedgehog depends on cilia architecture: *Dev Cell*, v. 12, p. 767-78.
- Cavenagh, M. M., M. Breiner, A. Schurmann, A. G. Rosenwald, T. Terui, C. Zhang, P. A. Randazzo, M. Adams, H. G. Joost, and R. A. Kahn, 1994, ADP-ribosylation factor (ARF)-like 3, a new member of the ARF family of GTP-binding proteins cloned from human and rat tissues: *J Biol Chem*, v. 269, p. 18937-42.
- Cevik, S., Y. Hori, O. I. Kaplan, K. Kida, T. Toivenon, C. Foley-Fisher, D. Cottell, T. Katada, K. Kontani, and O. E. Blacque, 2010, Joubert syndrome Arl13b functions at ciliary membranes and stabilizes protein transport in *Caenorhabditis elegans*: *J Cell Biol*, v. 188, p. 953-69.
- Cevik, S., A. A. Sanders, E. Van Wijk, K. Boldt, L. Clarke, J. van Reeuwijk, Y. Hori, N. Horn, L. Hetterschijt, A. Wdowicz, A. Mullins, K. Kida, O. I. Kaplan, S. E. van Beersum, K. Man Wu, S. J. Letteboer, D. A. Mans, T. Katada, K. Kontani, M. Ueffing, R. Roepman, H. Kremer, and O. E. Blacque, 2013, Active transport and diffusion barriers restrict Joubert Syndrome-associated ARL13B/ARL-13 to an Inv-like ciliary membrane subdomain: *PLoS Genet*, v. 9, p. e1003977.
- Chang, B., H. Khanna, N. Hawes, D. Jimeno, S. He, C. Lillo, S. K. Parapuram, H. Cheng, A. Scott, R. E. Hurd, J. A. Sayer, E. A. Otto, M. Attanasio, J. F. O'Toole, G. Jin, C. Shou, F. Hildebrandt, D. S. Williams, J. R. Heckenlively, and A. Swaroop, 2006, In-frame deletion in a novel centrosomal/ciliary protein CEP290/NPHP6 perturbs its interaction with RPGR and results in early-onset retinal degeneration in the rd16 mouse: *Hum Mol Genet*, v. 15, p. 1847-57.
- Chapple, J. P., A. J. Hardcastle, C. Grayson, L. A. Spackman, K. R. Willison, and M. E. Cheetham, 2000, Mutations in the N-terminus of the X-linked retinitis pigmentosa protein RP2 interfere with the normal targeting of the protein to the plasma membrane: *Hum Mol Genet*, v. 9, p. 1919-26.
- Chen, Z., J. Barbi, S. Bu, H. Y. Yang, Z. Li, Y. Gao, D. Jinasena, J. Fu, F. Lin, C. Chen, J. Zhang, N. Yu, X. Li, Z. Shan, J. Nie, Z. Gao, H. Tian, Y. Li, Z. Yao, Y. Zheng, B. V. Park, Z. Pan, E. Dang, H. Wang, W. Luo, L. Li, G. L. Semenza, S. G. Zheng, K. Loser, A. Tsun, M. I. Greene, D. M. Pardoll, F. Pan, and B. Li, 2013, The ubiquitin ligase Stub1 negatively modulates regulatory T cell suppressive activity by promoting degradation of the transcription factor Foxp3: *Immunity*, v. 39, p. 272-85.
- Chih, B., P. Liu, Y. Chinn, C. Chalouni, L. G. Komuves, P. E. Hass, W. Sandoval, and A. S. Peterson, 2012, A ciliopathy complex at the transition zone protects the cilia as a privileged membrane domain.: *Nat Cell Biol*, v. 14, p. 61-72.
- Christensen, S. T., S. F. Pedersen, P. Satir, I. R. Veland, and L. Schneider, 2008, The primary cilium coordinates signaling pathways in cell cycle control and migration during development and tissue repair.: *Curr Top Dev Biol*, v. 85, p. 261-301.
- Chávez, M., S. Ena, J. Van Sande, A. de Kerchove d'Exaerde, S. Schurmans, and S. N. Schiffmann, 2015, Modulation of Ciliary Phosphoinositide Content Regulates Trafficking and Sonic Hedgehog Signaling Output: *Dev Cell*, v. 34, p. 338-50.

- Clark, J., L. Moore, A. Krasinskas, J. Way, J. Battey, J. Tamkun, and R. A. Kahn, 1993, Selective amplification of additional members of the ADP-ribosylation factor (ARF) family: cloning of additional human and *Drosophila* ARF-like genes: *Proc Natl Acad Sci U S A*, v. 90, p. 8952-6.
- Cole, D. G., 2003, The intraflagellar transport machinery of *Chlamydomonas reinhardtii*: *Traffic*, v. 4, p. 435-42.
- Cole, D. G., D. R. Diener, A. L. Himelblau, P. L. Beech, J. C. Fuster, and J. L. Rosenbaum, 1998, *Chlamydomonas* kinesin-II-dependent intraflagellar transport (IFT): IFT particles contain proteins required for ciliary assembly in *Caenorhabditis elegans* sensory neurons: *J Cell Biol*, v. 141, p. 993-1008.
- Collins, B. M., P. J. Watson, and D. J. Owen, 2003, The structure of the GGA1-GAT domain reveals the molecular basis for ARF binding and membrane association of GGAs: *Dev Cell*, v. 4, p. 321-32.
- Corbit, K. C., P. Aanstad, V. Singla, A. R. Norman, D. Y. Stainier, and J. F. Reiter, 2005, Vertebrate Smoothed functions at the primary cilium: *Nature*, v. 437, p. 1018-21.
- Corbit, K. C., A. E. Shyer, W. E. Dowdle, J. Gaulden, V. Singla, M. H. Chen, P. T. Chuang, and J. F. Reiter, 2008, Kif3a constrains beta-catenin-dependent Wnt signalling through dual ciliary and non-ciliary mechanisms: *Nat Cell Biol*, v. 10, p. 70-6.
- Cortellino, S., C. Wang, B. Wang, M. R. Bassi, E. Caretti, D. Champeval, A. Calmont, M. Jarnik, J. Burch, K. S. Zaret, L. Larue, and A. Bellacosa, 2009, Defective ciliogenesis, embryonic lethality and severe impairment of the Sonic Hedgehog pathway caused by inactivation of the mouse complex A intraflagellar transport gene *Ift122/Wdr10*, partially overlapping with the DNA repair gene *Med1/Mbd4*: *Dev Biol*, v. 325, p. 225-37.
- Cox, J., and M. Mann, 2008, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification: *Nat Biotechnol*, v. 26, p. 1367-72.
- Cox, J., N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen, and M. Mann, 2011, Andromeda: a peptide search engine integrated into the MaxQuant environment: *J Proteome Res*, v. 10, p. 1794-805.
- Craige, B., C. C. Tsao, D. R. Diener, Y. Hou, K. F. Lechtreck, J. L. Rosenbaum, and G. B. Witman, 2010, CEP290 tethers flagellar transition zone microtubules to the membrane and regulates flagellar protein content.: *J Cell Biol*, v. 190, p. 927-40.
- Cuvillier, A., F. Redon, J. C. Antoine, P. Chardin, T. DeVos, and G. Merlin, 2000, LdARL-3A, a *Leishmania* promastigote-specific ADP-ribosylation factor-like protein, is essential for flagellum integrity: *J Cell Sci*, v. 113 ( Pt 11), p. 2065-74.
- Czarnecki, P. G., G. C. Gabriel, D. K. Manning, M. Sergeev, K. Lemke, N. T. Klena, X. Liu, Y. Chen, Y. Li, J. T. San Agustin, M. K. Garnaas, R. J. Francis, K. Tobita, W. Goessling, G. J. Pazour, C. W. Lo, D. R. Beier, and J. V. Shah, 2015, ANKS6 is the critical activator of NEK8 kinase in embryonic situs determination and organ patterning: *Nat Commun*, v. 6, p. 6023.
- Czarnecki, P. G., and J. V. Shah, 2012, The ciliary transition zone: from morphology and molecules to medicine.: *Trends Cell Biol*, v. 22, p. 201-10.
- Dang, H. Q., Q. Zhou, V. W. Rowlett, H. Hu, K. J. Lee, W. Margolin, and Z. Li, 2017, Proximity Interactions among Basal Body Components in *Trypanosoma brucei* Identify Novel Regulators of Basal Body Biogenesis and Inheritance: *MBio*, v. 8.
- Davidge, J. A., E. Chambers, H. A. Dickinson, K. Towers, M. L. Ginger, P. G. McKean, and K. Gull, 2006, Trypanosome IFT mutants provide insight into the motor location for mobility of the flagella connector and flagellar membrane formation: *J Cell Sci*, v. 119, p. 3935-43.
- Dawson, S. C., and S. A. House, 2010, Life with eight flagella: flagellar assembly and division in *Giardia*: *Curr Opin Microbiol*, v. 13, p. 480-90.

- Daugaard, M., M. Jäättelä, and M. Rohde, 2005, Hsp70-2 is required for tumor cell growth and survival: *Cell Cycle*, v. 4, p. 877-80.
- Daugaard, M., M. Rohde, and M. Jäättelä, 2007, The heat shock protein 70 family: Highly homologous proteins with overlapping and distinct functions: *FEBS Lett*, v. 581, p. 3702-10.
- Dean, S., F. Moreira-Leite, V. Varga, and K. Gull, 2016, Cilium transition zone proteome reveals compartmentalization and differential dynamics of ciliopathy complexes: *Proc Natl Acad Sci U S A*, v. 113, p. E5135-43.
- Dean, S., J. Sunter, R. J. Wheeler, I. Hodgkinson, E. Gluenz, and K. Gull, 2015, A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids: *Open Biol*, v. 5, p. 140197.
- Deane, J. A., D. G. Cole, E. S. Seeley, D. R. Diener, and J. L. Rosenbaum, 2001, Localization of intraflagellar transport protein IFT52 identifies basal body transitional fibers as the docking site for IFT particles: *Curr Biol*, v. 11, p. 1586-90.
- DeBonis, S., E. Neumann, and D. A. Skoufias, 2015, Self protein-protein interactions are involved in TPPP/p25 mediated microtubule bundling: *Sci Rep*, v. 5, p. 13242.
- De Craene, J. O., R. Ripp, O. Lecompte, J. D. Thompson, O. Poch, and S. Friant, 2012, Evolutionary analysis of the ENTH/ANTH/VHS protein superfamily reveals a coevolution between membrane trafficking and metabolism: *BMC Genomics*, v. 13, p. 297.
- Dekker, C., P. C. Stirling, E. A. McCormack, H. Filmore, A. Paul, R. L. Brost, M. Costanzo, C. Boone, M. R. Leroux, and K. R. Willison, 2008, The interaction network of the chaperonin CCT: *EMBO J*, v. 27, p. 1827-39.
- Delestré, L., Z. Bakey, C. Prado, S. Hoffmann, M. T. Bihoreau, B. Lelongt, and D. Gauguier, 2015, ANKS3 Co-Localises with ANKS6 in Mouse Renal Cilia and Is Associated with Vasopressin Signaling and Apoptosis In Vivo in Mice: *PLoS One*, v. 10, p. e0136781.
- Delous, M., L. Baala, R. Salomon, C. Laclef, J. Vierkotten, K. Tory, C. Golzio, T. Lacoste, L. Besse, C. Ozilou, I. Moutkine, N. E. Hellman, I. Anselme, F. Silbermann, C. Vesque, C. Gerhardt, E. Rattenberry, M. T. Wolf, M. C. Gubler, J. Martinovic, F. Encha-Razavi, N. Boddaert, M. Gonzales, M. A. Macher, H. Nivet, G. Champion, J. P. Berthélémy, P. Niaudet, F. McDonald, F. Hildebrandt, C. A. Johnson, M. Vekemans, C. Antignac, U. Rüther, S. Schneider-Maunoury, T. Attié-Bitach, and S. Saunier, 2007, The ciliary gene RPGRIP1L is mutated in cerebello-oculo-renal syndrome (Joubert syndrome type B) and Meckel syndrome: *Nat Genet*, v. 39, p. 875-81.
- Diaz, M. L., A. Solari, and C. I. Gonzalez, 2011, Differential expression of *Trypanosoma cruzi* I associated with clinical forms of Chagas disease: overexpression of oxidative stress proteins in acute patient isolate: *J Proteomics*, v. 74, p. 1673-82.
- Ding, X. F., C. Luo, K. Q. Ren, J. Zhang, J. L. Zhou, X. Hu, R. S. Liu, Y. Wang, and X. Gao, 2008, Characterization and expression of a human KCTD1 gene containing the BTB domain, which mediates transcriptional repression and homomeric interactions.: *DNA Cell Biol*, v. 27, p. 257-65.
- Dix, D. J., J. W. Allen, B. W. Collins, C. Mori, N. Nakamura, P. Poorman-Allen, E. H. Goulding, and E. M. Eddy, 1996, Targeted gene disruption of Hsp70-2 results in failed meiosis, germ cell apoptosis, and male infertility: *Proc Natl Acad Sci U S A*, v. 93, p. 3264-8.
- Dix, D. J., J. W. Allen, B. W. Collins, P. Poorman-Allen, C. Mori, D. R. Blizard, P. R. Brown, E. H. Goulding, B. D. Strong, and E. M. Eddy, 1997, HSP70-2 is required for desynapsis of synaptonemal complexes during meiotic prophase in juvenile and adult mouse spermatocytes: *Development*, v. 124, p. 4595-603.
- Dosztányi, Z., V. Csizmok, P. Tompa, and I. Simon, 2005, IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content: *Bioinformatics*, v. 21, p. 3433-4.

- Dowdle, W. E., J. F. Robinson, A. Kneist, M. S. Sirerol-Piquer, S. G. Frints, K. C. Corbit, N. A. Zaghloul, G. van Lijnschoten, L. Mulders, D. E. Verver, K. Zerres, R. R. Reed, T. Attié-Bitach, C. A. Johnson, J. M. García-Verdugo, N. Katsanis, C. Bergmann, and J. F. Reiter, 2011, Disruption of a ciliary B9 protein complex causes Meckel syndrome.: *Am J Hum Genet*, v. 89, p. 94-110.
- Drini, S., A. Criscuolo, P. Lechat, H. Imamura, T. Skalický, N. Rachidi, J. Lukeš, J. C. Dujardin, and G. F. Späth, 2016, Species- and Strain-Specific Adaptation of the HSP70 Super Family in Pathogenic Trypanosomatids: *Genome Biol Evol*, v. 8, p. 1980-95.
- Dudek, J., J. Benedix, S. Cappel, M. Greiner, C. Jalal, L. Müller, and R. Zimmermann, 2009, Functions and pathologies of BiP and its interaction partners: *Cell Mol Life Sci*, v. 66, p. 1556-69.
- Dunker, A. K., C. J. Brown, J. D. Lawson, L. M. Iakoucheva, and Z. Obradović, 2002, Intrinsic disorder and protein function: *Biochemistry*, v. 41, p. 6573-82.
- Dworniczak, B., and M. E. Mirault, 1987, Structure and expression of a human gene coding for a 71 kd heat shock 'cognate' protein: *Nucleic Acids Res*, v. 15, p. 5181-97.
- Easton, D. P., Y. Kaneko, and J. R. Subjeck, 2000, The hsp110 and Grp1 70 stress proteins: newly recognized relatives of the Hsp70s: *Cell Stress Chaperones*, v. 5, p. 276-90.
- Eguether, T., J. T. San Agustin, B. T. Keady, J. A. Jonassen, Y. Liang, R. Francis, K. Tobita, C. A. Johnson, Z. A. Abdelhamed, C. W. Lo, and G. J. Pazour, 2014, IFT27 links the BBSome to IFT for maintenance of the ciliary signaling compartment: *Dev Cell*, v. 31, p. 279-90.
- El-Sayed, N. M., P. J. Myler, G. Blandin, M. Berriman, J. Crabtree, G. Aggarwal, E. Caler, H. Renauld, E. A. Worthey, C. Hertz-Fowler, E. Ghedin, C. Peacock, D. C. Bartholomeu, B. J. Haas, A. N. Tran, J. R. Wortman, U. C. Alsmark, S. Angiuoli, A. Anupama, J. Badger, F. Bringaud, E. Cadag, J. M. Carlton, G. C. Cerqueira, T. Creasy, A. L. Delcher, A. Djikeng, T. M. Embley, C. Hauser, A. C. Ivens, S. K. Kummerfeld, J. B. Pereira-Leal, D. Nilsson, J. Peterson, S. L. Salzberg, J. Shallom, J. C. Silva, J. Sundaram, S. Westenberger, O. White, S. E. Melville, J. E. Donelson, B. Andersson, K. D. Stuart, and N. Hall, 2005, Comparative genomics of trypanosomatid parasitic protozoa.: *Science*, v. 309, p. 404-9.
- Engman, D. M., K. Henkle-Dührsen, L. V. Kirchhoff, and J. E. Donelson, 1995, *Trypanosoma cruzi*: accumulation of polycistronic hsp70 RNAs during severe heat shock: *Exp Parasitol*, v. 80, p. 575-7.
- Engstler, M., T. Pfohl, S. Herminghaus, M. Boshart, G. Wiegertjes, N. Heddergott, and P. Overath, 2007, Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes: *Cell*, v. 131, p. 505-15.
- Evans, C. G., L. Chang, and J. E. Gestwicki, 2010, Heat shock protein 70 (hsp70) as an emerging drug target: *J Med Chem*, v. 53, p. 4585-602.
- Evans, J. E., J. J. Snow, A. L. Gunnarson, G. Ou, H. Stahlberg, K. L. McDonald, and J. M. Scholey, 2006a, Functional modulation of IFT kinesins extends the sensory repertoire of ciliated neurons in *Caenorhabditis elegans*: *J Cell Biol*, v. 172, p. 663-9.
- Evans, R. J., A. J. Hardcastle, and M. E. Cheetham, 2006b, Focus on molecules: X-linked Retinitis Pigmentosa 2 protein, RP2: *Exp Eye Res*, v. 82, p. 543-4.
- Evans, R. J., N. Schwarz, K. Nagel-Wolfrum, U. Wolfrum, A. J. Hardcastle, and M. E. Cheetham, 2010, The retinitis pigmentosa protein RP2 links pericentriolar vesicle transport between the Golgi and the primary cilium: *Hum Mol Genet*, v. 19, p. 1358-67.
- Fan, Z. C., R. H. Behal, S. Geimer, Z. Wang, S. M. Williamson, H. Zhang, D. G. Cole, and H. Qin, 2010, *Chlamydomonas* IFT70/CrDYF-1 is a core component of IFT particle complex B and is required for flagellar assembly: *Mol Biol Cell*, v. 21, p. 2696-706.
- Feistel, K., and M. Blum, 2006, Three types of cilia including a novel 9+4 axoneme on the notochordal plate of the rabbit embryo: *Dev Dyn*, v. 235, p. 3348-58.



- Feldman, J. L., and W. F. Marshall, 2009, ASQ2 encodes a TBCC-like protein required for mother-daughter centriole linkage and mitotic spindle orientation: *Curr Biol*, v. 19, p. 1238-43.
- Field, M. C., and M. Carrington, 2009, The trypanosome flagellar pocket.: *Nat Rev Microbiol*, v. 7, p. 775-86.
- Field, M. C., T. Sergeenko, Y. N. Wang, S. Böhm, and M. Carrington, 2010, Chaperone requirements for biosynthesis of the trypanosome variant surface glycoprotein: *PLoS One*, v. 5, p. e8468.
- Flechas, I. D., A. Cuellar, Z. M. Cucunuba, F. Rosas, V. Velasco, M. Steindel, C. Thomas Mdel, M. C. Lopez, J. M. Gonzalez, and C. J. Puerta, 2009, Characterising the KMP11 and HSP-70 recombinant antigens' humoral immune response profile in chagasic patients: *BMC Infect Dis*, v. 9, p. 186.
- Flegontov, P., J. Votýpka, T. Skalický, M. D. Logacheva, A. A. Penin, G. Tanifuji, N. T. Onodera, A. S. Kondrashov, P. Volf, J. M. Archibald, and J. Lukeš, 2013, Paratrypanosoma is a novel early-branching trypanosomatid: *Curr Biol*, v. 23, p. 1787-93.
- Fliegauf, M., T. Benzing, and H. Omran, 2007, When cilia go bad: cilia defects and ciliopathies.: *Nat Rev Mol Cell Biol*, v. 8, p. 880-93.
- Fliegauf, M., A. Kahle, K. Häffner, and B. Zieger, 2014, Distinct localization of septin proteins to ciliary sub-compartments in airway epithelial cells: *Biol Chem*, v. 395, p. 151-6.
- Florin, L., K. A. Becker, C. Sapp, C. Lambert, H. Sirma, M. Müller, R. E. Streeck, and M. Sapp, 2004, Nuclear translocation of papillomavirus minor capsid protein L2 requires Hsc70: *J Virol*, v. 78, p. 5546-53.
- Folgueira, C., and J. M. Requena, 2007, A postgenomic view of the heat shock proteins in kinetoplastids: *FEMS Microbiol Rev*, v. 31, p. 359-77.
- Follit, J. A., L. Li, Y. Vucica, and G. J. Pazour, 2010, The cytoplasmic tail of fibrocystin contains a ciliary targeting sequence.: *J Cell Biol*, v. 188, p. 21-8.
- Ford, M. G., B. M. Pearse, M. K. Higgins, Y. Vallis, D. J. Owen, A. Gibson, C. R. Hopkins, P. R. Evans, and H. T. McMahon, 2001, Simultaneous binding of PtdIns(4,5)P2 and clathrin by AP180 in the nucleation of clathrin lattices on membranes: *Science*, v. 291, p. 1051-5.
- Fort, C., S. Bonnefoy, L. Kohl, and P. Bastin, 2016, Intraflagellar transport is required for the maintenance of the trypanosome flagellum composition but not its length: *J Cell Sci*, v. 129, p. 3026-41.
- Frevert, U., A. Movila, O. V. Nikolskaia, J. Raper, Z. B. Mackey, M. Abdulla, J. McKerrow, and D. J. Grab, 2012, Early invasion of brain parenchyma by African trypanosomes: *PLoS One*, v. 7, p. e43913.
- Gabernet-Castello, C., J. B. Dacks, and M. C. Field, 2009, The single ENTH-domain protein of trypanosomes; endocytic functions and evolutionary relationship with epsin: *Traffic*, v. 10, p. 894-911.
- Gao, C., and Y. G. Chen, 2010, Dishevelled: The hub of Wnt signaling: *Cell Signal*, v. 22, p. 717-27.
- Garcia-Gonzalo, F. R., K. C. Corbit, M. S. Sirerol-Piquer, G. Ramaswami, E. A. Otto, T. R. Noriega, A. D. Seol, J. F. Robinson, C. L. Bennett, D. J. Josifova, J. M. García-Verdugo, N. Katsanis, F. Hildebrandt, and J. F. Reiter, 2011, A transition zone complex regulates mammalian ciliogenesis and ciliary membrane composition.: *Nat Genet*, v. 43, p. 776-84.
- Garcia-Gonzalo, F. R., S. C. Phua, E. C. Roberson, G. Garcia, M. Abedin, S. Schurmans, T. Inoue, and J. F. Reiter, 2015, Phosphoinositides Regulate Ciliary Protein Trafficking to Modulate Hedgehog Signaling: *Dev Cell*, v. 34, p. 400-9.
- Garcia-Gonzalo, F. R., and J. F. Reiter, 2012, Scoring a backstage pass: Mechanisms of ciliogenesis and ciliary access.: *J Cell Biol*, v. 197, p. 697-709.

- Garcia-Gonzalo, F. R., and J. F. Reiter, 2017, Open Sesame: How Transition Fibers and the Transition Zone Control Ciliary Composition: *Cold Spring Harb Perspect Biol*, v. 9.
- Garcia-Mayoral, M. F., R. Castano, M. L. Fanarraga, J. C. Zabala, M. Rico, and M. Bruix, 2011, The solution structure of the N-terminal domain of human tubulin binding cofactor C reveals a platform for tubulin interaction: *PLoS One*, v. 6, p. e25912.
- Gerdes, J. M., E. E. Davis, and N. Katsanis, 2009, The vertebrate primary cilium in development, homeostasis, and disease.: *Cell*, v. 137, p. 32-45.
- Gerdes, J. M., Y. Liu, N. A. Zaghoul, C. C. Leitch, S. S. Lawson, M. Kato, P. A. Beachy, P. L. Beales, G. N. DeMartino, S. Fisher, J. L. Badano, and N. Katsanis, 2007, Disruption of the basal body compromises proteasomal function and perturbs intracellular Wnt response: *Nat Genet*, v. 39, p. 1350-60.
- Gilula, N. B., and P. Satir, 1972, The ciliary necklace. A ciliary membrane specialization: *J Cell Biol*, v. 53, p. 494-509.
- Ginger, M. L., 2014, Protein moonlighting in parasitic protists: *Biochem Soc Trans*, v. 42, p. 1734-9.
- Ginger, M. L., P. W. Collingridge, R. W. Brown, R. Sproat, M. K. Shaw, and K. Gull, 2013, Calmodulin is required for paraflagellar rod assembly and flagellum-cell body attachment in trypanosomes: *Protist*, v. 164, p. 528-40.
- Ginger, M. L., N. Portman, and P. G. McKean, 2008, Swimming with protists: perception, motility and flagellum assembly.: *Nat Rev Microbiol*, v. 6, p. 838-50.
- Glass, D. J., R. I. Polvere, and L. H. Van der Ploeg, 1986, Conserved sequences and transcription of the hsp70 gene family in *Trypanosoma brucei*: *Mol Cell Biol*, v. 6, p. 4657-66.
- Gluenz, E., M. L. Povelones, P. T. Englund, and K. Gull, 2011, The kinetoplast duplication cycle in *Trypanosoma brucei* is orchestrated by cytoskeleton-mediated cell morphogenesis: *Mol Cell Biol*, v. 31, p. 1012-21.
- Goate, A. M., D. N. Cooper, C. Hall, T. K. Leung, E. Solomon, and L. Lim, 1987, Localization of a human heat-shock HSP 70 gene sequence to chromosome 6 and detection of two other loci by somatic-cell hybrid and restriction fragment length polymorphism analysis: *Hum Genet*, v. 75, p. 123-8.
- Goetz, S. C., and K. V. Anderson, 2010, The primary cilium: a signalling centre during vertebrate development: *Nat Rev Genet*, v. 11, p. 331-44.
- Goetz, S. C., F. Bangs, C. L. Barrington, N. Katsanis, and K. V. Anderson, 2017, The Meckel syndrome- associated protein MKS1 functionally interacts with components of the BBSome and IFT complexes to mediate ciliary trafficking and hedgehog signaling: *PLoS One*, v. 12, p. e0173399.
- Goetz, S. C., K. F. Liem, and K. V. Anderson, 2012, The spinocerebellar ataxia-associated gene Tau tubulin kinase 2 controls the initiation of ciliogenesis: *Cell*, v. 151, p. 847-58.
- Goloubinoff, P., 2017, Editorial: The HSP70 Molecular Chaperone Machines: *Front Mol Biosci*, v. 4, p. 1.
- Gonçalves, J., S. Nolasco, R. Nascimento, M. Lopez Fanarraga, J. C. Zabala, and H. Soares, 2010, TBCCD1, a new centrosomal protein, is required for centrosome and Golgi apparatus positioning: *EMBO Rep*, v. 11, p. 194-200.
- Gonçalves, J., and L. Pelletier, 2017, The Ciliary Transition Zone: Finding the Pieces and Assembling the Gate: *Mol Cells*, v. 40, p. 243-253.
- Gorden, N. T., H. H. Arts, M. A. Parisi, K. L. Coene, S. J. Letteboer, S. E. van Beersum, D. A. Mans, A. Hikida, M. Eckert, D. Knutzen, A. F. Alswaid, H. Ozyurek, S. Dibooglu, E. A. Otto, Y. Liu, E. E. Davis, C. M. Hutter, T. K. Bammler, F. M. Farin, M. Dorschner, M. Topçu, E. H. Zackai, P. Rosenthal, K. N. Owens, N. Katsanis, J. B. Vincent, F. Hildebrandt, E. W. Rubel, D. W. Raible, N. V. Knoers, P. F. Chance, R. Roepman, C. B. Moens, I. A. Glass, and D. Doherty, 2008, CC2D2A is mutated in Joubert syndrome and interacts

- with the ciliopathy-associated basal body protein CEP290.: *Am J Hum Genet*, v. 83, p. 559-71.
- Gotthardt, K., M. Lokaj, C. Koerner, N. Falk, A. Gießl, and A. Wittinghofer, 2015, A G-protein activation cascade from Arl13B to Arl3 and implications for ciliary targeting of lipidated proteins: *Elife*, v. 4.
- Goujon, M., H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern, and R. Lopez, 2010, A new bioinformatics analysis tools framework at EMBL-EBI.: *Nucleic Acids Res*, v. 38, p. W695-9.
- Graser, S., Y. D. Stierhof, S. B. Lavoie, O. S. Gassner, S. Lamla, M. Le Clech, and E. A. Nigg, 2007, Cep164, a novel centriole appendage protein required for primary cilium formation: *J Cell Biol*, v. 179, p. 321-30.
- Grayson, C., F. Bartolini, J. P. Chapple, K. R. Willison, A. Bhamidipati, S. A. Lewis, P. J. Luthert, A. J. Hardcastle, N. J. Cowan, and M. E. Cheetham, 2002, Localization in the human retina of the X-linked retinitis pigmentosa protein RP2, its homologue cofactor C and the RP2 interacting protein Arl3: *Hum Mol Genet*, v. 11, p. 3065-74.
- Gruhler, A., J. V. Olsen, S. Mohammed, P. Mortensen, N. J. Faergeman, M. Mann, and O. N. Jensen, 2005, Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway: *Mol Cell Proteomics*, v. 4, p. 310-27.
- Gull, K., 1999, The cytoskeleton of trypanosomatid parasites: *Annu Rev Microbiol*, v. 53, p. 629-55.
- Guo, J., G. Jin, L. Meng, H. Ma, D. Nie, J. Wu, L. Yuan, and C. Shou, 2004, Subcellular localization of tumor-associated antigen 3H11Ag: *Biochem Biophys Res Commun*, v. 324, p. 922-30.
- Gupta-Rossi, N., S. Ortica, V. Meas-Yedid, S. Heuss, J. Moretti, J. C. Olivo-Marin, and A. Israël, 2011, The adaptor-associated kinase 1, AAK1, is a positive regulator of the Notch pathway: *J Biol Chem*, v. 286, p. 18720-30.
- Guther, M. L., M. D. Urbaniak, A. Tavendale, A. Prescott, and M. A. Ferguson, 2014, High-confidence glycosome proteome for procyclic form *Trypanosoma brucei* by epitope-tag organelle enrichment and SILAC proteomics: *J Proteome Res*, v. 13, p. 2796-806.
- Hammarton, T. C., 2007, Cell cycle regulation in *Trypanosoma brucei*.: *Mol Biochem Parasitol*, v. 153, p. 1-8.
- Han, Y. G., and A. Alvarez-Buylla, 2010, Role of primary cilia in brain development and cancer: *Curr Opin Neurobiol*, v. 20, p. 58-67.
- Hanks, S. K., A. M. Quinn, and T. Hunter, 1988, The protein kinase family: conserved features and deduced phylogeny of the catalytic domains: *Science*, v. 241, p. 42-52.
- Haycraft, C. J., B. Banizs, Y. Aydin-Son, Q. Zhang, E. J. Michaud, and B. K. Yoder, 2005, Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function: *PLoS Genet*, v. 1, p. e53.
- Hayes, P., V. Varga, S. Olego-Fernandez, J. Sunter, M. L. Ginger, and K. Gull, 2014, Modulation of a cytoskeletal calpain-like protein induces major transitions in trypanosome morphology: *J Cell Biol*, v. 206, p. 377-84.
- He, C. Y., M. Pypaert, and G. Warren, 2005, Golgi duplication in *Trypanosoma brucei* requires Centrin2: *Science*, v. 310, p. 1196-8.
- He, M., R. Subramanian, F. Bangs, T. Omelchenko, K. F. Liem, T. M. Kapoor, and K. V. Anderson, 2014, The kinesin-4 protein Kif7 regulates mammalian Hedgehog signalling by organizing the cilium tip compartment: *Nat Cell Biol*, v. 16, p. 663-72.
- Hill, K. L., 2003, Biology and mechanism of trypanosome cell motility: *Eukaryot Cell*, v. 2, p. 200-8.
- Hodges, M. E., N. Scheumann, B. Wickstead, J. A. Langdale, and K. Gull, 2010, Reconstructing the evolutionary history of the centriole from protein components: *J Cell Sci*, v. 123, p. 1407-13.

- Hoefele, J., R. Sudbrak, R. Reinhardt, S. Lehrack, S. Hennig, A. Imm, U. Muerb, B. Utsch, M. Attanasio, J. F. O'Toole, E. Otto, and F. Hildebrandt, 2005, Mutational analysis of the NPHP4 gene in 250 patients with nephronophthisis: *Hum Mutat*, v. 25, p. 411.
- Hoff, S., J. Halbritter, D. Epting, V. Frank, T. M. Nguyen, J. van Reeuwijk, C. Boehlke, C. Schell, T. Yasunaga, M. Helmstädter, M. Mergen, E. Filhol, K. Boldt, N. Horn, M. Ueffing, E. A. Otto, T. Eisenberger, M. W. Elting, J. A. van Wijk, D. Bockenhauer, N. J. Sebire, S. Rittig, M. Vyberg, T. Ring, M. Pohl, L. Pape, T. J. Neuhaus, N. A. Elshakhs, S. J. Koon, P. C. Harris, F. Grahammer, T. B. Huber, E. W. Kuehn, A. Kramer-Zucker, H. J. Bolz, R. Roepman, S. Saunier, G. Walz, F. Hildebrandt, C. Bergmann, and S. S. Lienkamp, 2013, ANKS6 is a central component of a nephronophthisis module linking NEK8 to INVS and NPHP3: *Nat Genet*, v. 45, p. 951-6.
- Hofmann, K., and L. Falquet, 2001, A ubiquitin-interacting motif conserved in components of the proteasomal and lysosomal protein degradation systems: *Trends Biochem Sci*, v. 26, p. 347-50.
- Holopainen, J. M., C. L. Cheng, L. L. Molday, G. Johal, J. Coleman, F. Dyka, T. Hii, J. Ahn, and R. S. Molday, 2010, Interaction and localization of the retinitis pigmentosa protein RP2 and NSF in retinal photoreceptor cells: *Biochemistry*, v. 49, p. 7439-47.
- Hu, H., Q. Zhou, and Z. Li, 2015, SAS-4 Protein in *Trypanosoma brucei* Controls Life Cycle Transitions by Modulating the Length of the Flagellum Attachment Zone Filament: *J Biol Chem*, v. 290, p. 30453-63.
- Hu, Q., and W. J. Nelson, 2011, Ciliary diffusion barrier: the gatekeeper for the primary cilium compartment: *Cytoskeleton (Hoboken)*, v. 68, p. 313-24.
- Hu, Q., W. J. Nelson, and E. T. Spiliotis, 2008, Forchlorfenuron alters mammalian septin assembly, organization, and dynamics: *J Biol Chem*, v. 283, p. 29563-71.
- Huang, L., N. F. Mivechi, and D. Moskopidhis, 2001, Insights into regulation and function of the major stress-induced hsp70 molecular chaperone in vivo: analysis of mice with targeted gene disruption of the hsp70.1 or hsp70.3 gene: *Mol Cell Biol*, v. 21, p. 8575-91.
- Huang, L., K. Szymanska, V. L. Jensen, A. R. Janecke, A. M. Innes, E. E. Davis, P. Frosk, C. Li, J. R. Willer, B. N. Chodirker, C. R. Greenberg, D. R. McLeod, F. P. Bernier, A. E. Chudley, T. Müller, M. Shboul, C. V. Logan, C. M. Loucks, C. L. Beaulieu, R. V. Bowie, S. M. Bell, J. Adkins, F. I. Zuniga, K. D. Ross, J. Wang, M. R. Ban, C. Becker, P. Nürnberg, S. Douglas, C. M. Craft, M. A. Akimenko, R. A. Hegele, C. Ober, G. Utermann, H. J. Bolz, D. E. Bulman, N. Katsanis, O. E. Blacque, D. Doherty, J. S. Parboosingh, M. R. Leroux, C. A. Johnson, and K. M. Boycott, 2011, TMEM237 is mutated in individuals with a Joubert syndrome related disorder and expands the role of the TMEM family at the ciliary transition zone.: *Am J Hum Genet*, v. 89, p. 713-30.
- Huangfu, D., and K. V. Anderson, 2005, Cilia and Hedgehog responsiveness in the mouse: *Proc Natl Acad Sci U S A*, v. 102, p. 11325-30.
- Huangfu, D., A. Liu, A. S. Rakeman, N. S. Murcia, L. Niswander, and K. V. Anderson, 2003, Hedgehog signalling in the mouse requires intraflagellar transport proteins: *Nature*, v. 426, p. 83-7.
- Hughes, L. C., K. S. Ralston, K. L. Hill, and Z. H. Zhou, 2012, Three-dimensional structure of the Trypanosome flagellum suggests that the paraflagellar rod functions as a biomechanical spring: *PLoS One*, v. 7, p. e25700
- Hutchings, N. R., J. E. Donelson, and K. L. Hill, 2002, Trypanin is a cytoskeletal linker protein and is required for cell motility in African trypanosomes.: *J Cell Biol*, v. 156, p. 867-77.
- Hyman, J., H. Chen, P. P. Di Fiore, P. De Camilli, and A. T. Brunger, 2000, Epsin 1 undergoes nucleocytoplasmic shuttling and its eps15 interactor NH(2)-terminal homology (ENTH) domain, structurally similar to Armadillo and HEAT repeats, interacts with the

- transcription factor promyelocytic leukemia Zn(2)+ finger protein (PLZF): *J Cell Biol*, v. 149, p. 537-46
- Ihara, M., A. Kinoshita, S. Yamada, H. Tanaka, A. Tanigaki, A. Kitano, M. Goto, K. Okubo, H. Nishiyama, O. Ogawa, C. Takahashi, S. Itohara, Y. Nishimune, M. Noda, and M. Kinoshita, 2005, Cortical organization by the septin cytoskeleton is essential for structural and mechanical integrity of mammalian spermatozoa: *Dev Cell*, v. 8, p. 343-52.
- Imamura, H., T. Downing, F. Van den Broeck, M. J. Sanders, S. Rijal, S. Sundar, A. Mannaert, M. Vanaerschot, M. Berg, G. De Muylder, F. Dumetz, B. Cuypers, I. Maes, M. Domagalska, S. Decuypere, K. Rai, S. Uranw, N. R. Bhattarai, B. Khanal, V. K. Prajapati, S. Sharma, O. Stark, G. Schöni, H. P. De Koning, L. Settimo, B. Vanhollebeke, S. Roy, B. Ostyn, M. Boelaert, L. Maes, M. Berriman, J. C. Dujardin, and J. A. Cotton, 2016, Evolutionary genomics of epidemic visceral leishmaniasis in the Indian subcontinent: *Elife*, v. 5.
- Iomini, C., V. Babaev-Khaimov, M. Sassaroli, and G. Piperno, 2001, Protein particles in *Chlamydomonas* flagella undergo a transport cycle consisting of four phases: *J Cell Biol*, v. 153, p. 13-24.
- Iomini, C., L. Li, J. M. Esparza, and S. K. Dutcher, 2009, Retrograde intraflagellar transport mutants identify complex A proteins with multiple genetic interactions in *Chlamydomonas reinhardtii*: *Genetics*, v. 183, p. 885-96.
- Ishiba, Y., T. Higashide, N. Mori, A. Kobayashi, S. Kubota, M. J. McLaren, H. Satoh, F. Wong, and G. Inana, 2007, Targeted inactivation of synaptic HRG4 (UNC119) causes dysfunction in the distal photoreceptor and slow retinal degeneration, revealing a new function: *Exp Eye Res*, v. 84, p. 473-85.
- Ishikawa, H., and W. F. Marshall, 2011, Ciliogenesis: building the cell's antenna.: *Nat Rev Mol Cell Biol*, v. 12, p. 222-34.
- Ismail, S. A., Y. X. Chen, A. Rusinova, A. Chandra, M. Bierbaum, L. Gremer, G. Triola, H. Waldmann, P. I. Bastiaens, and A. Wittinghofer, 2011, Arl2-GTP and Arl3-GTP regulate a GDI-like transport system for farnesylated cargo: *Nat Chem Biol*, v. 7, p. 942-9.
- Itoh, T., and P. De Camilli, 2006, BAR, F-BAR (EFC) and ENTH/ANTH domains in the regulation of membrane-cytosol interfaces and membrane curvature: *Biochim Biophys Acta*, v. 1761, p. 897-912.
- Jensen, V. L., C. Li, R. V. Bowie, L. Clarke, S. Mohan, O. E. Blacque, and M. R. Leroux, 2015, Formation of the transition zone by Mks5/Rpgrip1L establishes a ciliary zone of exclusion (CIZE) that compartmentalises ciliary signalling proteins and controls PIP2 ciliary abundance: *EMBO J*, v. 34, p. 2537-56.
- Jiang, H., and A. M. English, 2002, Quantitative analysis of the yeast proteome by incorporation of isotopically labeled leucine: *J Proteome Res*, v. 1, p. 345-50.
- Jiang, S. T., Y. Y. Chiou, E. Wang, Y. L. Chien, H. H. Ho, F. J. Tsai, C. Y. Lin, S. P. Tsai, and H. Li, 2009, Essential role of nephrocystin in photoreceptor intraflagellar transport in mouse: *Hum Mol Genet*, v. 18, p. 1566-77.
- Jiang, S. T., Y. Y. Chiou, E. Wang, H. K. Lin, S. P. Lee, H. Y. Lu, C. K. Wang, M. J. Tang, and H. Li, 2008, Targeted disruption of Nphp1 causes male infertility due to defects in the later steps of sperm morphogenesis in mice: *Hum Mol Genet*, v. 17, p. 3368-79.
- Jin, H., S. R. White, T. Shida, S. Schulz, M. Aguiar, S. P. Gygi, J. F. Bazan, and M. V. Nachury, 2010, The conserved Bardet-Biedl syndrome proteins assemble a coat that traffics membrane proteins to cilia: *Cell*, v. 141, p. 1208-19
- Jones, A., A. Faldas, A. Foucher, E. Hunt, A. Tait, J. M. Wastling, and C. M. Turner, 2006, Visualisation and analysis of proteomic data from the procyclic form of *Trypanosoma brucei*: *Proteomics*, v. 6, p. 259-67.

- Jones AJ, Grkovic T, Sykes ML, and Avery VM. 2013, Trypanocidal activity of marine natural products: *Mar Drugs*, 11(10):4058-82
- Jones, N. G., E. B. Thomas, E. Brown, N. J. Dickens, T. C. Hammarton, and J. C. Mottram, 2014, Regulators of *Trypanosoma brucei* cell cycle progression and differentiation identified using a kinome-wide RNAi screen: *PLoS Pathog*, v. 10, p. e1003886.
- Jovic, M., M. Sharma, J. Rahajeng, and S. Caplan, 2010, The early endosome: a busy sorting station for proteins at the crossroads: *Histol Histopathol*, v. 25, p. 99-112
- Kabututu, Z. P., M. Thayer, J. H. Melehani, and K. L. Hill, 2010, CMF70 is a subunit of the dynein regulatory complex: *J Cell Sci*, v. 123, p. 3587-95.
- Kabani, M., and C. N. Martineau, 2008, Multiple hsp70 isoforms in the eukaryotic cytosol: mere redundancy or functional specificity?: *Curr Genomics*, v. 9, p. 338-248.
- Kakihara, Y., and W. A. Houry, 2012, The R2TP complex: discovery and functions: *Biochim Biophys Acta*, v. 1823, p. 101-7.
- Kampinga, H. H., J. Hageman, M. J. Vos, H. Kubota, R. M. Tanguay, E. A. Bruford, M. E. Cheetham, B. Chen, and L. E. Hightower, 2009, Guidelines for the nomenclature of the human heat shock proteins: *Cell Stress Chaperones*, v. 14, p. 105-11.
- Katsanis, N., P. L. Beales, M. O. Woods, R. A. Lewis, J. S. Green, P. S. Parfrey, S. J. Ansley, W. S. Davidson, and J. R. Lupski, 2000, Mutations in MKKS cause obesity, retinal dystrophy and renal malformations associated with Bardet-Biedl syndrome: *Nat Genet*, v. 26, p. 67-70.
- Kay, B. K., M. Yamabhai, B. Wendland, and S. D. Emr, 1999, Identification of a novel domain shared by putative components of the endocytic and cytoskeletal machinery: *Protein Sci*, v. 8, p. 435-8.
- Keady, B. T., R. Samtani, K. Tobita, M. Tsuchya, J. T. San Agustin, J. A. Follit, J. A. Jonassen, R. Subramanian, C. W. Lo, and G. J. Pazour, 2012, IFT25 links the signal-dependent movement of Hedgehog components to intraflagellar transport: *Dev Cell*, v. 22, p. 940-51.
- Kelly, S., J. Reed, S. Kramer, L. Ellis, H. Webb, J. Sunter, J. Salje, N. Marinsek, K. Gull, B. Wickstead, and M. Carrington, 2007, Functional genomics in *Trypanosoma brucei*: a collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci.: *Mol Biochem Parasitol*, v. 154, p. 103-9.
- Kerner, M. J., D. J. Naylor, Y. Ishihama, T. Maier, H. C. Chang, A. P. Stines, C. Georgopoulos, D. Frishman, M. Hayer-Hartl, M. Mann, and F. U. Hartl, 2005, Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*: *Cell*, v. 122, p. 209-20.
- Kiang, J. G., and G. C. Tsokos, 1998, Heat shock protein 70 kDa: molecular biology, biochemistry, and physiology: *Pharmacol Ther*, v. 80, p. 183-201.
- Kim, J., S. R. Krishnaswami, and J. G. Gleeson, 2008, CEP290 interacts with the centriolar satellite component PCM-1 and is required for Rab8 localization to the primary cilium: *Hum Mol Genet*, v. 17, p. 3796-805.
- Kim, S. K., A. Shindo, T. J. Park, E. C. Oh, S. Ghosh, R. S. Gray, R. A. Lewis, C. A. Johnson, T. Attie-Bittach, N. Katsanis, and J. B. Wallingford, 2010, Planar cell polarity acts through septins to control collective cell movement and ciliogenesis: *Science*, v. 329, p. 1337-40.
- Kim, D. I., S. C. Jensen, K. A. Noble, B. Kc, K. H. Roux, K. Motamedchaboki, and K. J. Roux, 2016, An improved smaller biotin ligase for BioID proximity labeling: *Mol Biol Cell*, v. 27, p. 1188-96.
- King, N., M. J. Westbrook, S. L. Young, A. Kuo, M. Abedin, J. Chapman, S. Fairclough, U. Hellsten, Y. Isogai, I. Letunic, M. Marr, D. Pincus, N. Putnam, A. Rokas, K. J. Wright, R. Zuzow, W. Dirks, M. Good, D. Goodstein, D. Lemons, W. Li, J. B. Lyons, A. Morris, S. Nichols, D. J. Richter, A. Salamov, J. G. Sequencing, P. Bork, W. A. Lim, G. Manning, W. T. Miller, W. McGinnis, H. Shapiro, R. Tjian, I. V. Grigoriev, and D. Rokhsar, 2008, The

- genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans: *Nature*, v. 451, p. 783-8.
- King, S. M., 2016, Axonemal Dynein Arms: *Cold Spring Harb Perspect Biol*, v. 8.
- Kissel, H., M. M. Georgescu, S. Larisch, K. Manova, G. R. Hunnicutt, and H. Steller, 2005, The Sept4 septin locus is required for sperm terminal differentiation in mice: *Dev Cell*, v. 8, p. 353-64.
- Knighton, D. R., J. H. Zheng, L. F. Ten Eyck, V. A. Ashford, N. H. Xuong, S. S. Taylor, and J. M. Sowadski, 1991, Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase: *Science*, v. 253, p. 407-14.
- Kobayashi, T., S. Kim, Y. C. Lin, T. Inoue, and B. D. Dynlacht, 2014, The CP110-interacting proteins Talpid3 and Cep290 play overlapping and distinct roles in cilia assembly: *J Cell Biol*, v. 204, p. 215-29.
- Kohl, L., D. Robinson, and P. Bastin, 2003b, [The flagellum: from cell motility to morphogenesis]. *J Soc Biol*, v. 197, p. 379-87.
- Kohl, L., D. Robinson, and P. Bastin, 2003, Novel roles for the flagellum in cell morphogenesis and cytokinesis of trypanosomes.: *EMBO J*, v. 22, p. 5336-46.
- Kovács, G. G., L. László, J. Kovács, P. H. Jensen, E. Lindersson, G. Botond, T. Molnár, A. Perczel, F. Hudecz, G. Mezo, A. Erdei, L. Tirián, A. Lehotzky, E. Gelpi, H. Budka, and J. Ovádi, 2004, Natively unfolded tubulin polymerization promoting protein TPPP/p25 is a common marker of alpha-synucleinopathies: *Neurobiol Dis*, v. 17, p. 155-62.
- Koyfman, A. Y., M. F. Schmid, L. Gheiratmand, C. J. Fu, H. A. Khant, D. Huang, C. Y. He, and W. Chiu, 2011, Structure of *Trypanosoma brucei* flagellum accounts for its bihelical motion: *Proc Natl Acad Sci U S A*, v. 108, p. 11105-8.
- Kozminski, K. G., P. L. Beech, and J. L. Rosenbaum, 1995, The *Chlamydomonas* kinesin-like protein FLA10 is involved in motility associated with the flagellar membrane: *J Cell Biol*, v. 131, p. 1517-27.
- Kozminski, K. G., K. A. Johnson, P. Forscher, and J. L. Rosenbaum, 1993, A motility in the eukaryotic flagellum unrelated to flagellar beating: *Proc Natl Acad Sci U S A*, v. 90, p. 5519-23.
- Kubota, H., G. Hynes, and K. Willison, 1995, The chaperonin containing t-complex polypeptide 1 (TCP-1). Multisubunit machinery assisting in protein folding and assembly in the eukaryotic cytosol: *Eur J Biochem*, v. 230, p. 3-16.
- Kuhnel, K., S. Veltel, I. Schlichting, and A. Wittinghofer, 2006, Crystal structure of the human retinitis pigmentosa 2 protein and its interaction with Arl3: *Structure*, v. 14, p. 367-78.
- Kuriyama, R., C. Besse, M. Gèze, C. K. Omoto, and J. Schrével, 2005, Dynamic organization of microtubules and microtubule-organizing centers during the sexual phase of a parasitic protozoan, *Lecudina tuzetae* (Gregarine, Apicomplexa): *Cell Motil Cytoskeleton*, v. 62, p. 195-209.
- Käser, S., S. Oeljeklaus, J. Týč, S. Vaughan, B. Warscheid, and A. Schneider, 2016, Outer membrane protein functions as integrator of protein import and DNA inheritance in mitochondria: *Proc Natl Acad Sci U S A*, v. 113, p. E4467-75.
- Lacomble, S., S. Vaughan, C. Gadelha, M. K. Morphew, M. K. Shaw, J. R. McIntosh, and K. Gull, 2010, Basal body movements orchestrate membrane organelle division and cell morphogenesis in *Trypanosoma brucei*: *J Cell Sci*, v. 123, p. 2884-91.
- LaCount, D. J., B. Barrett, and J. E. Donelson, 2002, *Trypanosoma brucei* FLA1 is required for flagellum attachment and cytokinesis: *J Biol Chem*, v. 277, p. 17580-8.
- Lai, C. K., N. Gupta, X. Wen, L. Rangell, B. Chih, A. S. Peterson, J. F. Bazan, L. Li, and S. J. Scales, 2011, Functional characterization of putative cilia genes by high-content analysis: *Mol Biol Cell*, v. 22, p. 1104-19.
- Lambacher, N. J., A. L. Bruel, T. J. van Dam, K. Szymańska, G. G. Slaats, S. Kuhns, G. J. McManus, J. E. Kennedy, K. Gaff, K. M. Wu, R. van der Lee, L. Burglen, D. Doummar, J. B. Rivière,

- L. Faivre, T. Attié-Bitach, S. Saunier, A. Curd, M. Peckham, R. H. Giles, C. A. Johnson, M. A. Huynen, C. Thauvin-Robinet, and O. E. Blacque, 2016, TMEM107 recruits ciliopathy proteins to subdomains of the ciliary transition zone and causes Joubert syndrome: *Nat Cell Biol*, v. 18, p. 122-31.
- Langousis, G., and K. L. Hill, 2014, Motility and more: the flagellum of *Trypanosoma brucei*: *Nat Rev Microbiol*, v. 12, p. 505-18.
- Lehtreck, K. F., J. M. Brown, J. L. Sampaio, J. M. Craft, A. Shevchenko, J. E. Evans, and G. B. Witman, 2013, Cycling of the signaling protein phospholipase D through cilia requires the BBSome only for the export phase: *J Cell Biol*, v. 201, p. 249-61.
- Lee, S. H., M. Kim, B. W. Yoon, Y. J. Kim, S. J. Ma, J. K. Roh, J. S. Lee, and J. S. Seo, 2001, Targeted hsp70.1 disruption increases infarction volume after focal cerebral ischemia in mice: *Stroke*, v. 32, p. 2905-12.
- Leettola, C. N., M. J. Knight, D. Cascio, S. Hoffman, and J. U. Bowie, 2014, Characterization of the SAM domain of the PKD-related protein ANKS6 and its interaction with ANKS3: *BMC Struct Biol*, v. 14, p. 17.
- Lewis, S. A., G. Tian, and N. J. Cowan, 1997, The alpha- and beta-tubulin folding pathways: *Trends Cell Biol*, v. 7, p. 479-84.
- Lehotzky, A., P. Lau, N. Tokési, N. Muja, L. D. Hudson, and J. Ovádi, 2010, Tubulin polymerization-promoting protein (TPPP/p25) is critical for oligodendrocyte differentiation: *Glia*, v. 58, p. 157-68.
- Li, C., V. L. Jensen, K. Park, J. Kennedy, F. R. Garcia-Gonzalo, M. Romani, R. De Mori, A. L. Bruel, D. Gaillard, B. Doray, E. Lopez, J. B. Rivière, L. Faivre, C. Thauvin-Robinet, J. F. Reiter, O. E. Blacque, E. M. Valente, and M. R. Leroux, 2016, MKS5 and CEP290 Dependent Assembly Pathway of the Ciliary Transition Zone: *PLoS Biol*, v. 14, p. e1002416.
- Li, L., N. Khan, T. Hurd, A. K. Ghosh, C. Cheng, R. Molday, J. R. Heckenlively, A. Swaroop, and H. Khanna, 2013, Ablation of the X-linked retinitis pigmentosa 2 (Rp2) gene in mice results in opsin mislocalization and photoreceptor degeneration: *Invest Ophthalmol Vis Sci*, v. 54, p. 4503-11.
- Li, Y., and J. Hu, 2011, Small GTPases and cilia: *Protein Cell*, v. 2, p. 13-25.
- Li, Y., K. Ling, and J. Hu, 2012, The emerging role of Arf/Arl small GTPases in cilia and ciliopathies: *J Cell Biochem*, v. 113, p. 2201-7.
- Li, Z., and C. C. Wang, 2008, KMP-11, a basal body and flagellar protein, is required for cell division in *Trypanosoma brucei*: *Eukaryot Cell*, v. 7, p. 1941-50.
- Liem, K. F., A. Ashe, M. He, P. Satir, J. Moran, D. Beier, C. Wicking, and K. V. Anderson, 2012, The IFT-A complex regulates Shh signaling through cilia structure and membrane protein trafficking: *J Cell Biol*, v. 197, p. 789-800.
- Lindersson, E., D. Lundvig, C. Petersen, P. Madsen, J. R. Nyengaard, P. Højrup, T. Moos, D. Otzen, W. P. Gai, P. C. Blumbergs, and P. H. Jensen, 2005, p25alpha Stimulates alpha-synuclein aggregation and is co-localized with aggregated alpha-synuclein in alpha-synucleinopathies: *J Biol Chem*, v. 280, p. 5703-15.
- Liu, F., J. Chen, S. Yu, R. K. Raghupathy, X. Liu, Y. Qin, C. Li, M. Huang, S. Liao, J. Wang, J. Zou, X. Shu, Z. Tang, and M. Liu, 2015, Knockout of RP2 decreases GRK1 and rod transducin subunits and leads to photoreceptor degeneration in zebrafish: *Hum Mol Genet*, v. 24, p. 4648-59.
- Llorca, O., J. Martín-Benito, J. Grantham, M. Ritco-Vonsovici, K. R. Willison, J. L. Carrascosa, and J. M. Valpuesta, 2001, The 'sequential allosteric ring' mechanism in the eukaryotic chaperonin-assisted folding of actin and tubulin: *EMBO J*, v. 20, p. 4065-75.
- Llorca, O., M. G. Smyth, S. Marco, J. L. Carrascosa, K. R. Willison, and J. M. Valpuesta, 1998, ATP binding induces large conformational changes in the apical and equatorial domains of the eukaryotic chaperonin containing TCP-1 complex: *J Biol Chem*, v. 273, p. 10091-4.



- Lohi, O., A. Poussu, Y. Mao, F. Quiocho, and V. P. Lehto, 2002, VHS domain -- a longshoreman of vesicle lines: *FEBS Lett*, v. 513, p. 19-23.
- Lokaj, M., S. K. Kösling, C. Koerner, S. M. Lange, S. E. van Beersum, J. van Reeuwijk, R. Roepman, N. Horn, M. Ueffing, K. Boldt, and A. Wittinghofer, 2015, The Interaction of CCDC104/BARTL1 with Arl3 and Implications for Ciliary Function: *Structure*, v. 23, p. 2122-32.
- Loktev, A. V., Q. Zhang, J. S. Beck, C. C. Searby, T. E. Scheetz, J. F. Bazan, D. C. Slusarski, V. C. Sheffield, P. K. Jackson, and M. V. Nachury, 2008, A BBSome subunit links ciliogenesis, microtubule stability, and acetylation: *Dev Cell*, v. 15, p. 854-65.
- Lu, Q., C. Insinna, C. Ott, J. Stauffer, P. A. Pintado, J. Rahajeng, U. Baxa, V. Walia, A. Cuenca, Y. S. Hwang, I. O. Daar, S. Lopes, J. Lippincott-Schwartz, P. K. Jackson, S. Caplan, and C. J. Westlake, 2015, Early steps in primary cilium assembly require EHD1/EHD3-dependent ciliary vesicle formation: *Nat Cell Biol*, v. 17, p. 531.
- Lucker, B. F., R. H. Behal, H. Qin, L. C. Siron, W. D. Taggart, J. L. Rosenbaum, and D. G. Cole, 2005, Characterization of the intraflagellar transport complex B core: direct interaction of the IFT81 and IFT74/72 subunits: *J Biol Chem*, v. 280, p. 27688-96.
- Lukes, J., D. L. Guilbride, J. Votýpka, A. Zíková, R. Benne, and P. T. Englund, 2002, Kinetoplast DNA network: evolution of an improbable structure: *Eukaryot Cell*, v. 1, p. 495-502.
- MacFarlane, J., M. L. Blaxter, R. P. Bishop, M. A. Miles, and J. M. Kelly, 1990, Identification and characterisation of a *Leishmania donovani* antigen belonging to the 70-kDa heat-shock protein family: *Eur J Biochem*, v. 190, p. 377-84.
- MacGregor, P., N. J. Savill, D. Hall, and K. R. Matthews, 2011, Transmission stages dominate trypanosome within-host dynamics during chronic infections: *Cell Host Microbe*, v. 9, p. 310-8.
- Maerker, T., E. van Wijk, N. Overlack, F. F. Kersten, J. McGee, T. Goldmann, E. Sehn, R. Roepman, E. J. Walsh, H. Kremer, and U. Wolfrum, 2008, A novel Usher protein network at the periciliary reloading point between molecular transport machineries in vertebrate photoreceptor cells: *Hum Mol Genet*, v. 17, p. 71-86.
- Maldonado-Báez, L., and B. Wendland, 2006, Endocytic adaptors: recruiters, coordinators and regulators: *Trends Cell Biol*, v. 16, p. 505-13.
- Manna, P. T., S. O. Obado, C. Boehm, C. Gadelha, A. Sali, B. T. Chait, M. P. Rout, and M. C. Field, 2017, Lineage-specific proteins essential for endocytosis in trypanosomes: *J Cell Sci*, v. 130, p. 1379-1392.
- Manning, G., G. D. Plowman, T. Hunter, and S. Sudarsanam, 2002, Evolution of protein kinase signaling from yeast to man: *Trends Biochem Sci*, v. 27, p. 514-20.
- Mao, Y., A. Nickitenko, X. Duan, T. E. Lloyd, M. N. Wu, H. Bellen, and F. A. Quiocho, 2000, Crystal structure of the VHS and FYVE tandem domains of Hrs, a protein involved in membrane trafficking and signal transduction: *Cell*, v. 100, p. 447-56.
- Mao, Y., J. Chen, J. A. Maynard, B. Zhang, and F. A. Quiocho, 2001, A novel all helix fold of the AP180 amino-terminal domain for phosphoinositide binding and clathrin assembly in synaptic vesicle endocytosis: *Cell*, v. 104, p. 433-40.
- Marcello, L., and J. D. Barry, 2007, From silent genes to noisy populations-dialogue between the genotype and phenotypes of antigenic variation: *J Eukaryot Microbiol*, v. 54, p. 14-7.
- Matthews, K. R., 2005, The developmental cell biology of *Trypanosoma brucei*: *J Cell Sci*, v. 118, p. 283-90.
- Matthews, K. R., J. R. Ellis, and A. Paterou, 2004, Molecular regulation of the life cycle of African trypanosomes: *Trends Parasitol*, v. 20, p. 40-7.
- Mayer, M. P., D. Brehmer, C. S. Gässler, and B. Bukau, 2001, Hsp70 chaperone machines: *Adv Protein Chem*, v. 59, p. 1-44.

- Mayer, M. P., and B. Bukau, 2005, Hsp70 chaperones: cellular functions and molecular mechanism: *Cell Mol Life Sci*, v. 62, p. 670-84.
- McAllaster, M. R., K. N. Ikeda, A. Lozano-Núñez, D. Anrather, V. Unterwurzacher, T. Gossenreiter, J. A. Perry, R. Crickley, C. J. Mercadante, S. Vaughan, and C. L. de Graffenried, 2015, Proteomic identification of novel cytoskeletal proteins associated with TbPLK, an essential regulator of cell morphogenesis in *Trypanosoma brucei*: *Mol Biol Cell*, v. 26, p. 3013-29.
- McKean, P. G., 2003, Coordination of cell cycle and cytokinesis in *Trypanosoma brucei*: *Curr Opin Microbiol*, v. 6, p. 600-7.
- McNicoll, F., J. Drummelsmith, M. Müller, E. Madore, N. Boilard, M. Ouellette, and B. Papadopoulou, 2006, A combined proteomic and transcriptomic approach to the study of stage differentiation in *Leishmania infantum*: *Proteomics*, v. 6, p. 3567-81.
- Michels, P. A., F. Bringaud, M. Herman, and V. Hannaert, 2006, Metabolic functions of glycosomes in trypanosomatids: *Biochim Biophys Acta*, v. 1763, p. 1463-77.
- Middelbeek, J., K. Clark, H. Venselaar, M. A. Huynen, and F. N. van Leeuwen, 2010, The alpha-kinase family: an exceptional branch on the protein kinase tree: *Cell Mol Life Sci*, v. 67, p. 875-90.
- Mill, P., P. J. Lockhart, E. Fitzpatrick, H. S. Mountford, E. A. Hall, M. A. Reijns, M. Keighren, M. Bahlo, C. J. Bromhead, P. Budd, S. Aftimos, M. B. Delatycki, R. Savarirayan, I. J. Jackson, and D. J. Amor, 2011, Human and mouse mutations in WDR35 cause short-rib polydactyly syndromes due to abnormal ciliogenesis: *Am J Hum Genet*, v. 88, p. 508-15.
- Min, W., F. Angileri, H. Luo, A. Lauria, M. Shanmugasundaram, A. M. Almerico, F. Cappello, E. C. de Macario, I. K. Lednev, A. J. Macario, and F. T. Robb, 2014, A human CCT5 gene mutation causing distal neuropathy impairs hexadecamer assembly in an archaeal model: *Sci Rep*, v. 4, p. 6688.
- Misra, S., B. M. Beach, and J. H. Hurley, 2000, Structure of the VHS domain of human Tom1 (target of myb 1): insights into interactions with proteins and membranes: *Biochemistry*, v. 39, p. 11282-90.
- Moran, J., P. McKean, and M. Ginger, 2014, Eukaryotic Flagella: Variations in Form, Function, and Composition during Evolution: *Bioscience*, v. 64, p. 12.
- Mori, R., and T. Toda, 2013, The dual role of fission yeast Tbc1/cofactor C orchestrates microtubule homeostasis in tubulin folding and acts as a GAP for GTPase Alp41/Arl2: *Mol Biol Cell*, v. 24, p. 1713-24, S1-8.
- Morriswood, B., K. Havlicek, L. Demmel, S. Yavuz, M. Sealey-Cardona, K. Vidilaseris, D. Anrather, J. Kostan, K. Djinovic-Carugo, K. J. Roux, and G. Warren, 2013, Novel bilobe components in *Trypanosoma brucei* identified using proximity-dependent biotinylation: *Eukaryot Cell*, v. 12, p. 356-67.
- Morriswood, B., C. Y. He, M. Sealey-Cardona, J. Yelinek, M. Pypaert, and G. Warren, 2009, The bilobe structure of *Trypanosoma brucei* contains a MORN-repeat protein: *Mol Biochem Parasitol*, v. 167, p. 95-103.
- Mueller, J., C. A. Perrone, R. Bower, D. G. Cole, and M. E. Porter, 2005, The FLA3 KAP subunit is required for localization of kinesin-2 to the site of flagellar assembly and processive anterograde intraflagellar transport: *Mol Biol Cell*, v. 16, p. 1341-54.
- Mykytyn, K., R. F. Mullins, M. Andrews, A. P. Chiang, R. E. Swiderski, B. Yang, T. Braun, T. Casavant, E. M. Stone, and V. C. Sheffield, 2004, Bardet-Biedl syndrome type 4 (BBS4)-null mice implicate Bbs4 in flagella formation but not global cilia assembly: *Proc Natl Acad Sci U S A*, v. 101, p. 8664-9.
- Nachury, M. V., A. V. Loktev, Q. Zhang, C. J. Westlake, J. Peränen, A. Merdes, D. C. Slusarski, R. H. Scheller, J. F. Bazan, V. C. Sheffield, and P. K. Jackson, 2007, A core complex of BBS

- proteins cooperates with the GTPase Rab8 to promote ciliary membrane biogenesis: *Cell*, v. 129, p. 1201-13.
- Nachury, M. V., E. S. Seeley, and H. Jin, 2010, Trafficking to the ciliary membrane: how to get across the periciliary diffusion barrier?: *Annu Rev Cell Dev Biol*, v. 26, p. 59-87.
- Naula, C., M. Parsons, and J. C. Mottram, 2005, Protein kinases as drug targets in trypanosomes and Leishmania: *Biochim Biophys Acta*, v. 1754, p. 151-9.
- Nett, I. R., D. M. Martin, D. Miranda-Saavedra, D. Lamont, J. D. Barber, A. Mehlert, and M. A. Ferguson, 2009, The phosphoproteome of bloodstream form *Trypanosoma brucei*, causative agent of African sleeping sickness: *Mol Cell Proteomics*, v. 8, p. 1527-38.
- Nicastro, D., X. Fu, T. Heuser, A. Tso, M. E. Porter, and R. W. Linck, 2011, Cryo-electron tomography reveals conserved features of doublet microtubules in flagella: *Proc Natl Acad Sci U S A*, v. 108, p. E845-53.
- Nirmalan, N., P. F. Sims, and J. E. Hyde, 2004, Quantitative proteomics of the human malaria parasite *Plasmodium falciparum* and its application to studies of development and inhibition: *Mol Microbiol*, v. 52, p. 1187-99.
- Nonaka, S., Y. Tanaka, Y. Okada, S. Takeda, A. Harada, Y. Kanai, M. Kido, and N. Hirokawa, 1998, Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein: *Cell*, v. 95, p. 829-37.
- Novarino, G., N. Akizu, and J. G. Gleeson, 2011, Modeling human disease in humans: the ciliopathies: *Cell*, v. 147, p. 70-9.
- Oberholzer, M., M. A. Lopez, B. T. McLelland, and K. L. Hill, 2010, Social motility in african trypanosomes: *PLoS Pathog*, v. 6, p. e1000739.
- Ogunuga, O. O., V. M. Longshaw, and G. L. Blatch, 2004, Hop: more than an Hsp70/Hsp90 adaptor protein: *Bioessays*, v. 26, p. 1058-68.
- Ogbadoyi, E. O., D. R. Robinson, and K. Gull, 2003, A high-order trans-membrane structural linkage is responsible for mitochondrial genome positioning and segregation by flagellar basal bodies in trypanosomes: *Mol Biol Cell*, v. 14, p. 1769-79.
- Olcese, C., M. P. Patel, A. Shoemark, S. Kiviluoto, M. Legendre, H. J. Williams, C. K. Vaughan, J. Hayward, A. Goldenberg, R. D. Emes, M. M. Munye, L. Dyer, T. Cahill, J. Bevilard, C. Gehrig, M. Guipponi, S. Chantot, P. Duquesnoy, L. Thomas, L. Jeanson, B. Copin, A. Tamalet, C. Thauvin-Robinet, J. F. Papon, A. Garin, I. Pin, G. Vera, P. Aurora, M. R. Fassad, L. Jenkins, C. Boustred, T. Cullup, M. Dixon, A. Onoufriadis, A. Bush, E. M. Chung, S. E. Antonarakis, M. R. Loebinger, R. Wilson, M. Armengot, E. Escudier, C. Hogg, S. Amselem, Z. Sun, L. Bartoloni, J. L. Blouin, H. M. Mitchison, and U. K. R. Group, 2017, X-linked primary ciliary dyskinesia due to mutations in the cytoplasmic axonemal dynein assembly factor PIH1D3: *Nat Commun*, v. 8, p. 14279.
- Omran, H., 2010, NPHP proteins: gatekeepers of the ciliary compartment: *J Cell Biol*, v. 190, p. 715-7.
- Omran, H., D. Kobayashi, H. Olbrich, T. Tsukahara, N. T. Loges, H. Hagiwara, Q. Zhang, G. Leblond, E. O'Toole, C. Hara, H. Mizuno, H. Kawano, M. Fliegauf, T. Yagi, S. Koshida, A. Miyawaki, H. Zentgraf, H. Seithe, R. Reinhardt, Y. Watanabe, R. Kamiya, D. R. Mitchell, and H. Takeda, 2008, Ktu/PF13 is required for cytoplasmic pre-assembly of axonemal dyneins: *Nature*, v. 456, p. 611-6.
- Ong, S. E., B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, and M. Mann, 2002, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics: *Mol Cell Proteomics*, v. 1, p. 376-86.
- Ong, S. E., and M. Mann, 2005, Mass spectrometry-based proteomics turns quantitative: *Nat Chem Biol*, v. 1, p. 252-62.
- Ong, S. E., and M. Mann, 2006, A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC): *Nat Protoc*, v. 1, p. 2650-60.

- Opperdoes, F. R., and P. Borst, 1977, Localization of nine glycolytic enzymes in a microbody-like organelle in *Trypanosoma brucei*: the glycosome: *FEBS Lett*, v. 80, p. 360-4.
- Orosz, F., G. G. Kovács, A. Lehotzky, J. Oláh, O. Vincze, and J. Ovádi, 2004, TPPP/p25: from unfolded protein to misfolding disease: prediction and experiments: *Biol Cell*, v. 96, p. 701-11.
- Orosz, F., and J. Ovádi, 2008, TPPP orthologs are ciliary proteins: *FEBS Lett*, v. 582, p. 3757-64.
- Orosz, F., 2012, A new protein superfamily: TPPP-like proteins: *PLoS One*, v. 7, p. e49276.
- Otto, E., J. Hoefele, R. Ruf, A. M. Mueller, K. S. Hiller, M. T. Wolf, M. J. Schuermann, A. Becker, R. Birkenhäger, R. Sudbrak, H. C. Hennies, P. Nürnberg, and F. Hildebrandt, 2002, A gene mutated in nephronophthisis and retinitis pigmentosa encodes a novel protein, nephroretinin, conserved in evolution: *Am J Hum Genet*, v. 71, p. 1161-7.
- Otto, E. A., B. Schermer, T. Obara, J. F. O'Toole, K. S. Hiller, A. M. Mueller, R. G. Ruf, J. Hoefele, F. Beekmann, D. Landau, J. W. Foreman, J. A. Goodship, T. Strachan, A. Kispert, M. T. Wolf, M. F. Gagnadoux, H. Nivet, C. Antignac, G. Walz, I. A. Drummond, T. Benzing, and F. Hildebrandt, 2003, Mutations in *INVS* encoding inversin cause nephronophthisis type 2, linking renal cystic disease to the function of primary cilia and left-right axis determination: *Nat Genet*, v. 34, p. 413-20.
- Otzen, D. E., D. M. Lundvig, R. Wimmer, L. H. Nielsen, J. R. Pedersen, and P. H. Jensen, 2005, p25alpha is flexible but natively folded and binds tubulin with oligomeric stoichiometry: *Protein Sci*, v. 14, p. 1396-409.
- Ou, G., O. E. Blacque, J. J. Snow, M. R. Leroux, and J. M. Scholey, 2005, Functional coordination of intraflagellar transport motors: *Nature*, v. 436, p. 583-7.
- Ovádi, J., F. Orosz, A. Lehotzky, and A. Lebotzky, 2005, What is the biological significance of the brain-specific tubulin-polymerization promoting protein (TPPP/p25)?: *IUBMB Life*, v. 57, p. 765-8.
- Pappenberger, G., J. A. Wilsher, S. M. Roe, D. J. Counsell, K. R. Willison, and L. H. Pearl, 2002, Crystal structure of the CCTgamma apical domain: implications for substrate binding to the eukaryotic cytosolic chaperonin: *J Mol Biol*, v. 318, p. 1367-79.
- Parsons, M., E. A. Worthey, P. N. Ward, and J. C. Mottram, 2005, Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*: *BMC Genomics*, v. 6, p. 127.
- Park, Z. Y., and D. H. Russell, 2001, Identification of individual proteins in complex protein mixtures by high-resolution, high-mass-accuracy MALDI TOF-mass spectrometry analysis of in-solution thermal denaturation/enzymatic digestion: *Anal Chem*, v. 73, p. 2558-64.
- Paulick, M. G., and C. R. Bertozzi, 2008, The glycosylphosphatidylinositol anchor: a complex membrane-anchoring structure for proteins: *Biochemistry*, v. 47, p. 6991-7000.
- Pazour, G. J., B. L. Dickert, and G. B. Witman, 1999, The DHC1b (DHC2) isoform of cytoplasmic dynein is required for flagellar assembly: *J Cell Biol*, v. 144, p. 473-81.
- Pazour, G. J., C. G. Wilkerson, and G. B. Witman, 1998, A dynein light chain is essential for the retrograde particle movement of intraflagellar transport (IFT): *J Cell Biol*, v. 141, p. 979-92.
- Pedersen, L. B., S. Geimer, and J. L. Rosenbaum, 2006, Dissecting the molecular mechanisms of intraflagellar transport in *Chlamydomonas*: *Curr Biol*, v. 16, p. 450-9.
- Pedersen, L. B., and J. L. Rosenbaum, 2008, Intraflagellar transport (IFT) role in ciliary assembly, resorption and signalling: *Curr Top Dev Biol*, v. 85, p. 23-61.
- Pelham, H. R., 1984, Hsp70 accelerates the recovery of nucleolar morphology after heat shock: *EMBO J*, v. 3, p. 3095-100.
- Pepin, J., and J. Donelson, 1999, African trypanosomiasis (sleeping sickness): *Tropical Infectious Diseases: Principles, Pathogens and Practice*, Churchill Livingstone, 774-84 p.

- Perdomo, D., M. Bonhivers, and D. R. Robinson, 2016, The Trypanosome Flagellar Pocket Collar and Its Ring Forming Protein-TbBILBO1: Cells, v. 5.
- Pigino, G., and T. Ishikawa, 2012, Axonemal radial spokes: 3D structure, function and assembly: Bioarchitecture, v. 2, p. 50-58.
- Piperno, G., and K. Mead, 1997, Transport of a novel complex in the cytoplasmic matrix of Chlamydomonas flagella: Proc Natl Acad Sci U S A, v. 94, p. 4457-62.
- Piperno, G., E. Siuda, S. Henderson, M. Segil, H. Vaananen, and M. Sassaroli, 1998, Distinct mutants of retrograde intraflagellar transport (IFT) share similar morphological and molecular defects: J Cell Biol, v. 143, p. 1591-601.
- Poon, S. K., L. Peacock, W. Gibson, K. Gull, and S. Kelly, 2012, A modular and optimized single marker system for generating Trypanosoma brucei cell lines expressing T7 RNA polymerase and the tetracycline repressor: Open Biol, v. 2, p. 110037.
- Portman, N., and K. Gull, 2010, The paraflagellar rod of kinetoplastid parasites: from structure to components and function: Int J Parasitol, v. 40, p. 135-48.
- Portman, N., and K. Gull, 2014, Identification of paralogous life-cycle stage specific cytoskeletal proteins in the parasite Trypanosoma brucei: PLoS One, v. 9, p. e106777.
- Poussu, A., O. Lohi, and V. P. Lehto, 2000, Vear, a novel Golgi-associated protein with VHS and gamma-adaptin "ear" domains: J Biol Chem, v. 275, p. 7176-83.
- Pradel, L. C., M. Bonhivers, N. Landrein, and D. R. Robinson, 2006, NIMA-related kinase TbNRKC is involved in basal body separation in Trypanosoma brucei: J Cell Sci, v. 119, p. 1852-63.
- Prodromou, N. V., C. L. Thompson, D. P. Osborn, K. F. Cogger, R. Ashworth, M. M. Knight, P. L. Beales, and J. P. Chapple, 2012, Heat shock induces rapid resorption of primary cilia: J Cell Sci, v. 125, p. 4297-305.
- Price, H. P., A. Peltan, M. Stark, and D. F. Smith, 2010, The small GTPase ARL2 is required for cytokinesis in Trypanosoma brucei: Mol Biochem Parasitol, v. 173, p. 123-31.
- Ralston, K. S., Z. P. Kabututu, J. H. Melehan, M. Oberholzer, and K. L. Hill, 2009, The Trypanosoma brucei flagellum: moving parasites in new directions.: Annu Rev Microbiol, v. 63, p. 335-62.
- Ralston, K. S., A. G. Lerner, D. R. Diener, and K. L. Hill, 2006, Flagellar motility contributes to cytokinesis in Trypanosoma brucei and is modulated by an evolutionarily conserved dynein regulatory system: Eukaryot Cell, v. 5, p. 696-711.
- Raviol, H., B. Bukau, and M. P. Mayer, 2006, Human and yeast Hsp110 chaperones exhibit functional differences: FEBS Lett, v. 580, p. 168-74.
- Reck, J., A. M. Schauer, K. VanderWaal Mills, R. Bower, D. Tritschler, C. A. Perrone, and M. E. Porter, 2016, The role of the dynein light intermediate chain in retrograde IFT and flagellar function in Chlamydomonas: Mol Biol Cell, v. 27, p. 2404-22.
- Redmond, S., J. Vadivelu, and M. C. Field, 2003, RNAi: an automated web-based tool for the selection of RNAi targets in Trypanosoma brucei: Mol Biochem Parasitol, v. 128, p. 115-8.
- Reese, T. S., 1965, OLFACTORY CILIA IN THE FROG: J Cell Biol, v. 25, p. 209-30.
- Reiter, J. F., O. E. Blacque, and M. R. Leroux, 2012, The base of the cilium: roles for transition fibres and the transition zone in ciliary formation, maintenance and compartmentalization.: EMBO Rep.
- Requena, J. M., A. Jimenez-Ruiz, M. Soto, R. Assiego, J. F. Santarén, M. C. Lopez, M. E. Patarroyo, and C. Alonso, 1992, Regulation of hsp70 expression in Trypanosoma cruzi by temperature and growth phase: Mol Biochem Parasitol, v. 53, p. 201-11.
- Richey, E. A., and H. Qin, 2012, Dissecting the sequential assembly and localization of intraflagellar transport particle complex B in Chlamydomonas: PLoS One, v. 7, p. e43118.

- Ridgley, E., P. Webster, C. Patton, and L. Ruben, 2000, Calmodulin-binding properties of the paraflagellar rod complex from *Trypanosoma brucei*: *Mol Biochem Parasitol*, v. 109, p. 195-201.
- Ringo, D. L., 1967, Flagellar motion and fine structure of the flagellar apparatus in *Chlamydomonas*: *J Cell Biol*, v. 33, p. 543-71.
- Rios, R. M., and M. Bornens, 2003, The Golgi apparatus at the cell centre: *Curr Opin Cell Biol*, v. 15, p. 60-6.
- Roberson, E. C., W. E. Dowdle, A. Ozanturk, F. R. Garcia-Gonzalo, C. Li, J. Halbritter, N. Elkhartoufi, J. D. Porath, H. Cope, A. Ashley-Koch, S. Gregory, S. Thomas, J. A. Sayer, S. Saunier, E. A. Otto, N. Katsanis, E. E. Davis, T. Attié-Bitach, F. Hildebrandt, M. R. Leroux, and J. F. Reiter, 2015, TMEM231, mutated in orofacioidigital and Meckel syndromes, organizes the ciliary transition zone: *J Cell Biol*, v. 209, p. 129-42.
- Robinson, D. R., and K. Gull, 1991, Basal body movements as a mechanism for mitochondrial genome segregation in the trypanosome cell cycle: *Nature*, v. 352, p. 731-3.
- Robinson, D. R., T. Sherwin, A. Ploubidou, E. H. Byard, and K. Gull, 1995, Microtubule polarity and dynamics in the control of organelle positioning, segregation, and cytokinesis in the trypanosome cell cycle: *J Cell Biol*, v. 128, p. 1163-72.
- Roepman, R., and U. Wolfrum, 2007, Protein networks and complexes in photoreceptor cilia: *Subcell Biochem*, v. 43, p. 209-35.
- Rohatgi, R., and W. J. Snell, 2010, The ciliary membrane: *Curr Opin Cell Biol*, v. 22, p. 541-6.
- Rosenbaum, J. L., and G. B. Witman, 2002, Intraflagellar transport.: *Nat Rev Mol Cell Biol*, v. 3, p. 813-25.
- Rosenfeld, J., J. Capdevielle, J. C. Guillemot, and P. Ferrara, 1992, In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis: *Anal Biochem*, v. 203, p. 173-9.
- Roux, K. J., D. I. Kim, M. Raida, and B. Burke, 2012, A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells: *J Cell Biol*, v. 196, p. 801-10.
- Rohde, M., M. Daugaard, M. H. Jensen, K. Helin, J. Nylandsted, and M. Jäättelä, 2005, Members of the heat-shock protein 70 family promote cancer cell growth by distinct mechanisms: *Genes Dev*, v. 19, p. 570-82.
- Rokas, A., D. Krüger, and S. B. Carroll, 2005, Animal evolution and the molecular signature of radiations compressed in time: *Science*, v. 310, p. 1933-8.
- Roseman, A. M., S. Chen, H. White, K. Braig, and H. R. Saibil, 1996, The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL: *Cell*, v. 87, p. 241-51.
- Rotureau, B., T. Blisnick, I. Subota, D. Julkowska, N. Cayet, S. Perrot, and P. Bastin, 2014, Flagellar adhesion in *Trypanosoma brucei* relies on interactions between different skeletal structures in the flagellum and cell body: *J Cell Sci*, v. 127, p. 204-15
- Saito, K., N. Tochio, T. Kigawa, and S. Yokoyama, 2007, Solution structure of the C-terminal region in human tubulin folding cofactor C, RCSB Protein data bank.
- Sahin, A., G. Lemercier, E. Tetaud, B. Espiau, P. Myler, K. Stuart, N. Bakalara, and G. Merlin, 2004, Trypanosomatid flagellum biogenesis: ARL-3A is involved in several species: *Exp Parasitol*, v. 108, p. 126-33.
- Sang, L., J. J. Miller, K. C. Corbit, R. H. Giles, M. J. Brauer, E. A. Otto, L. M. Baye, X. Wen, S. J. Scales, M. Kwong, E. G. Huntzicker, M. K. Sfakianos, W. Sandoval, J. F. Bazan, P. Kulkarni, F. R. Garcia-Gonzalo, A. D. Seol, J. F. O'Toole, S. Held, H. M. Reutter, W. S. Lane, M. A. Rafiq, A. Noor, M. Ansar, A. R. Devi, V. C. Sheffield, D. C. Slusarski, J. B. Vincent, D. A. Doherty, F. Hildebrandt, J. F. Reiter, and P. K. Jackson, 2011a, Mapping

- the NPHP-JBTS-MKS protein network reveals ciliopathy disease genes and pathways: *Cell*, v. 145, p. 513-28.
- Satir, P., 1968, Studies on cilia. 3. Further studies on the cilium tip and a "sliding filament" model of ciliary motility: *J Cell Biol*, v. 39, p. 77-94.
- Satir, P., and S. T. Christensen, 2007, Overview of structure and function of mammalian cilia.: *Annu Rev Physiol*, v. 69, p. 377-400.
- Satouh, Y., P. Padma, T. Toda, N. Satoh, H. Ide, and K. Inaba, 2005, Molecular characterization of radial spoke subcomplex containing radial spoke protein 3 and heat shock protein 40 in sperm flagella of the ascidian *Ciona intestinalis*: *Mol Biol Cell*, v. 16, p. 626-36.
- Sayer, J. A., E. A. Otto, J. F. O'Toole, G. Nurnberg, M. A. Kennedy, C. Becker, H. C. Hennies, J. Helou, M. Attanasio, B. V. Fausett, B. Utsch, H. Khanna, Y. Liu, I. Drummond, I. Kawakami, T. Kusakabe, M. Tsuda, L. Ma, H. Lee, R. G. Larson, S. J. Allen, C. J. Wilkinson, E. A. Nigg, C. Shou, C. Lillo, D. S. Williams, B. Hoppe, M. J. Kemper, T. Neuhaus, M. A. Parisi, I. A. Glass, M. Petry, A. Kispert, J. Gloy, A. Ganner, G. Walz, X. Zhu, D. Goldman, P. Nurnberg, A. Swaroop, M. R. Leroux, and F. Hildebrandt, 2006, The centrosomal protein nephrocystin-6 is mutated in Joubert syndrome and activates transcription factor ATF4: *Nat Genet*, v. 38, p. 674-81.
- Scheffzek, K., M. R. Ahmadian, W. Kabsch, L. Wiesmüller, A. Lautwein, F. Schmitz, and A. Wittinghofer, 1997, The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants: *Science*, v. 277, p. 333-8.
- Scheffzek, K., M. R. Ahmadian, and A. Wittinghofer, 1998, GTPase-activating proteins: helping hands to complement an active site: *Trends Biochem Sci*, v. 23, p. 257-62.
- Schlecht, R., A. H. Erbse, B. Bukau, and M. P. Mayer, 2011, Mechanics of Hsp70 chaperones enables differential interaction with client proteins: *Nat Struct Mol Biol*, v. 18, p. 345-51.
- Schmidt, K. N., S. Kuhns, A. Neuner, B. Hub, H. Zentgraf, and G. Pereira, 2012, Cep164 mediates vesicular docking to the mother centriole during early steps of ciliogenesis: *J Cell Biol*, v. 199, p. 1083-101.
- Schneider, A., A. Hemphill, T. Wyler, and T. Seebeck, 1988, Large microtubule-associated protein of *T. brucei* has tandemly repeated, near-identical sequences: *Science*, v. 241, p. 459-62.
- Schrack, J. J., P. Vogel, A. Abuin, B. Hampton, and D. S. Rice, 2006, ADP-ribosylation factor-like 3 is involved in kidney and photoreceptor development: *Am J Pathol*, v. 168, p. 1288-98.
- Schwahn, U., S. Lenzner, J. Dong, S. Feil, B. Hinzmann, G. van Duijnhoven, R. Kirschner, M. Hemberger, A. A. Bergen, T. Rosenberg, A. J. Pinckers, R. Fundele, A. Rosenthal, F. P. Cremers, H. H. Ropers, and W. Berger, 1998, Positional cloning of the gene for X-linked retinitis pigmentosa 2: *Nat Genet*, v. 19, p. 327-32.
- Schwarz, N., A. J. Hardcastle, and M. E. Cheetham, 2012a, Arl3 and RP2 mediated assembly and traffic of membrane associated cilia proteins: *Vision Res*, v. 75, p. 2-4.
- Schwarz, N., A. J. Hardcastle, and M. E. Cheetham, 2012b, The role of the X-linked retinitis pigmentosa protein RP2 in vesicle traffic and cilia function: *Adv Exp Med Biol*, v. 723, p. 527-32.
- Schwarz, N., A. Lane, K. Jovanovich, D. A. Parfitt, M. Aguila, C. L. Thompson, L. da Cruz, P. J. Coffey, J. P. Chapple, A. J. Hardcastle, and M. E. Cheetham, 2017, Arl3 and RP2 regulate the trafficking of ciliary tip kinesins: *Hum Mol Genet*.
- Schwarz, N., T. V. Novoselova, R. Wait, A. J. Hardcastle, and M. E. Cheetham, 2012c, The X-linked retinitis pigmentosa protein RP2 facilitates G protein traffic: *Hum Mol Genet*, v. 21, p. 863-73.
- Seeger-Nukpezah, T., J. L. Little, V. Serzhanova, and E. A. Golemis, 2013, Cilia and cilia-associated proteins in cancer: *Drug Discov Today Dis Mech*, v. 10, p. e135-e142.

- Seo, S., L. M. Baye, N. P. Schulz, J. S. Beck, Q. Zhang, D. C. Slusarski, and V. C. Sheffield, 2010, BBS6, BBS10, and BBS12 form a complex with CCT/TRiC family chaperonins and mediate BBSome assembly: *Proc Natl Acad Sci U S A*, v. 107, p. 1488-93.
- Sharer, J. D., and R. A. Kahn, 1999, The ARF-like 2 (ARL2)-binding protein, BART. Purification, cloning, and initial characterization: *J Biol Chem*, v. 274, p. 27553-61.
- Sharer, J. D., J. F. Shern, H. Van Valkenburgh, D. C. Wallace, and R. A. Kahn, 2002, ARL2 and BART enter mitochondria and bind the adenine nucleotide transporter: *Mol Biol Cell*, v. 13, p. 71-83.
- Shapiro, J., J. Ingram, and K. A. Johnson, 2005, Characterization of a molecular chaperone present in the eukaryotic flagellum: *Eukaryot Cell*, v. 4, p. 1591-4.
- Sherwin, T., and K. Gull, 1989a, The cell division cycle of *Trypanosoma brucei brucei*: timing of event markers and cytoskeletal modulations: *Philos Trans R Soc Lond B Biol Sci*, v. 323, p. 573-88.
- Sherwin, T., and K. Gull, 1989b, The cell division cycle of *Trypanosoma brucei brucei*: timing of event markers and cytoskeletal modulations.: *Philos Trans R Soc Lond B Biol Sci*, v. 323, p. 573-88.
- Shi, Y., D. D. Mosser, and R. I. Morimoto, 1998, Molecular chaperones as HSF1-specific transcriptional repressors: *Genes Dev*, v. 12, p. 654-66.
- Shiba, D., D. K. Manning, H. Koga, D. R. Beier, and T. Yokoyama, 2010, Inv acts as a molecular anchor for Nphp3 and Nek8 in the proximal segment of primary cilia: *Cytoskeleton (Hoboken)*, v. 67, p. 112-9.
- Shiba, T., M. Kawasaki, H. Takatsu, T. Nogi, N. Matsugaki, N. Igarashi, M. Suzuki, R. Kato, K. Nakayama, and S. Wakatsuki, 2003, Molecular mechanism of membrane recruitment of GGA by ARF in lysosomal protein transport: *Nat Struct Biol*, v. 10, p. 386-93.
- Shim, E. H., J. I. Kim, E. S. Bang, J. S. Heo, J. S. Lee, E. Y. Kim, J. E. Lee, W. Y. Park, S. H. Kim, H. S. Kim, O. Smithies, J. J. Jang, D. I. Jin, and J. S. Seo, 2002, Targeted disruption of hsp70.1 sensitizes to osmotic stress: *EMBO Rep*, v. 3, p. 857-61.
- Sillibourne, J. E., C. G. Specht, I. Izeddin, I. Hurbain, P. Tran, A. Triller, X. Darzacq, M. Dahan, and M. Bornens, 2011, Assessing the localization of centrosomal proteins by PALM/STORM nanoscopy: *Cytoskeleton (Hoboken)*, v. 68, p. 619-27.
- Sinden, R. E., and M. E. Smalley, 1976, Gametocytes of *Plasmodium falciparum*: phagocytosis by leucocytes in vivo and in vitro: *Trans R Soc Trop Med Hyg*, v. 70, p. 344-5.
- Skjoerringe, T., D. M. Lundvig, P. H. Jensen, and T. Moos, 2006, P25alpha/Tubulin polymerization promoting protein expression by myelinating oligodendrocytes of the developing rat brain: *J Neurochem*, v. 99, p. 333-42.
- Snow, J. J., G. Ou, A. L. Gunnarson, M. R. Walker, H. M. Zhou, I. Brust-Mascher, and J. M. Scholey, 2004, Two anterograde intraflagellar transport motors cooperate to build sensory cilia on *C. elegans* neurons: *Nat Cell Biol*, v. 6, p. 1109-13.
- Son, W. Y., C. T. Han, S. H. Hwang, J. H. Lee, S. Kim, and Y. C. Kim, 2000, Repression of hspA2 messenger RNA in human testes with abnormal spermatogenesis: *Fertil Steril*, v. 73, p. 1138-44.
- Sorokin, S., 1962, Centrioles and the formation of rudimentary cilia by fibroblasts and smooth muscle cells: *J Cell Biol*, v. 15, p. 363-77.
- Sorokin, S. P., 1968, Reconstructions of centriole formation and ciliogenesis in mammalian lungs: *J Cell Sci*, v. 3, p. 207-30.
- Spiegelman, B. M., S. M. Penningroth, and M. W. Kirschner, 1977, Turnover of tubulin and the N site GTP in Chinese hamster ovary cells: *Cell*, v. 12, p. 587-600.
- Spiess, C., A. S. Meyer, S. Reissmann, and J. Frydman, 2004, Mechanism of the eukaryotic chaperonin: protein folding in the chamber of secrets: *Trends Cell Biol*, v. 14, p. 598-604.



- Stebeck, C. E., R. P. Beecroft, B. N. Singh, A. Jardim, R. W. Olafson, C. Tuckey, K. D. Prenevost, and T. W. Pearson, 1995, Kinetoplastid membrane protein-11 (KMP-11) is differentially expressed during the life cycle of African trypanosomes and is found in a wide variety of kinetoplastid parasites: *Mol Biochem Parasitol*, v. 71, p. 1-13.
- Stephan, A., S. Vaughan, M. K. Shaw, K. Gull, and P. G. McKean, 2007, An essential quality control mechanism at the eukaryotic basal body prior to intraflagellar transport.: *Traffic*, v. 8, p. 1323-30.
- Stoetzel, C., J. Muller, V. Laurier, E. E. Davis, N. A. Zaghoul, S. Vicaire, C. Jacquelin, F. Plewniak, C. C. Leitch, P. Sarda, C. Hamel, T. J. de Ravel, R. A. Lewis, E. Friederich, C. Thibault, J. M. Danse, A. Verloes, D. Bonneau, N. Katsanis, O. Poch, J. L. Mandel, and H. Dollfus, 2007, Identification of a novel BBS gene (BBS12) highlights the major role of a vertebrate-specific branch of chaperonin-related proteins in Bardet-Biedl syndrome: *Am J Hum Genet*, v. 80, p. 1-11.
- Stoetzel, C., V. Laurier, E. E. Davis, J. Muller, S. Rix, J. L. Badano, C. C. Leitch, N. Salem, E. Chouery, S. Corbani, N. Jalk, S. Vicaire, P. Sarda, C. Hamel, D. Lacombe, M. Holder, S. Odent, S. Holder, A. S. Brooks, N. H. Elcioglu, E. D. Silva, E. Da Silva, B. Rossillion, S. Sigaudy, T. J. de Ravel, R. A. Lewis, B. Leheup, A. Verloes, P. Amati-Bonneau, A. Mégarbané, O. Poch, D. Bonneau, P. L. Beales, J. L. Mandel, N. Katsanis, and H. Dollfus, 2006, BBS10 encodes a vertebrate-specific chaperonin-like protein and is a major BBS locus: *Nat Genet*, v. 38, p. 521-4.
- Subota, I., D. Julkowska, L. Vincensini, N. Reeg, J. Buisson, T. Blisnick, D. Huet, S. Perrot, J. Santi-Rocca, M. Duchateau, V. Hourdel, J. C. Rousselle, N. Cayet, A. Namane, J. Chamot-Rooke, and P. Bastin, 2014, Proteomic analysis of intact flagella of procyclic *Trypanosoma brucei* cells identifies novel flagellar proteins with unique sub-localization and dynamics: *Mol Cell Proteomics*, v. 13, p. 1769-86.
- Sui, H., and K. H. Downing, 2006, Molecular architecture of axonemal microtubule doublets revealed by cryo-electron tomography: *Nature*, v. 442, p. 475-8.
- Sun, S. Y., C. Wang, Y. A. Yuan, and C. Y. He, 2013, An intracellular membrane junction consisting of flagellum adhesion glycoproteins links flagellum biogenesis to cell morphogenesis in *Trypanosoma brucei*: *J Cell Sci*, v. 126, p. 520-31.
- Sung, C. H., and M. R. Leroux, 2013, The roles of evolutionarily conserved functional modules in cilia-related trafficking: *Nat Cell Biol*, v. 15, p. 1387-97.
- Sunter, J. D., C. Benz, J. Andre, S. Whipple, P. G. McKean, K. Gull, M. L. Ginger, and J. Lukeš, 2015, Modulation of flagellum attachment zone protein FLAM3 and regulation of the cell shape in *Trypanosoma brucei* life cycle transitions: *J Cell Sci*, v. 128, p. 3117-30.
- Sunter, J. D., and K. Gull, 2016, The Flagellum Attachment Zone: 'The Cellular Ruler' of Trypanosome Morphology: *Trends Parasitol*, v. 32, p. 309-24.
- Szymanska, K., and C. A. Johnson, 2012, The transition zone: an essential functional compartment of cilia: *Cilia*, v. 1, p. 10.
- Sánchez, I., and B. D. Dynlacht, 2016, Cilium assembly and disassembly: *Nat Cell Biol*, v. 18, p. 711-7.
- Takahashi, M., K. Tomizawa, K. Sato, A. Ohtake, and A. Omori, 1995, A novel tau-tubulin kinase from bovine brain: *FEBS Lett*, v. 372, p. 59-64.
- Takahashi, M., K. Tomizawa, K. Ishiguro, K. Sato, A. Omori, S. Sato, A. Shiratsuchi, T. Uchida, and K. Imahori, 1991, A novel brain-specific 25 kDa protein (p25) is phosphorylated by a Ser/Thr-Pro kinase (TPK II) from tau protein kinase fractions: *FEBS Lett*, v. 289, p. 37-43.
- Takahashi, M., K. Tomizawa, S. C. Fujita, K. Sato, T. Uchida, and K. Imahori, 1993, A brain-specific protein p25 is localized and associated with oligodendrocytes, neuropil, and fiber-like structures of the CA3 hippocampal region in the rat brain: *J Neurochem*, v. 60, p. 228-35.

- Takaki, E., M. Fujimoto, T. Nakahari, S. Yonemura, Y. Miyata, N. Hayashida, K. Yamamoto, R. B. Vallee, T. Mikuriya, K. Sugahara, H. Yamashita, S. Inouye, and A. Nakai, 2007, Heat shock transcription factor 1 is required for maintenance of ciliary beating in mice: *J Biol Chem*, v. 282, p. 37285-92.
- Takata, H., M. Kato, K. Denda, and N. Kitamura, 2000, A hrs binding protein having a Src homology 3 domain is involved in intracellular degradation of growth factors and their receptors: *Genes Cells*, v. 5, p. 57-69.
- Takeshita, T., T. Arita, H. Asao, N. Tanaka, M. Higuchi, H. Kuroda, K. Kaneko, H. Munakata, Y. Endo, T. Fujita, and K. Sugamura, 1996, Cloning of a novel signal-transducing adaptor molecule containing an SH3 domain and ITAM: *Biochem Biophys Res Commun*, v. 225, p. 1035-9.
- Takeshita, T., T. Arita, M. Higuchi, H. Asao, K. Endo, H. Kuroda, N. Tanaka, K. Murata, N. Ishii, and K. Sugamura, 1997, STAM, signal transducing adaptor molecule, is associated with Janus kinases and involved in signaling for cell growth and c-myc induction: *Immunity*, v. 6, p. 449-57.
- Tamkun, J. W., R. A. Kahn, M. Kissinger, B. J. Brizuela, C. Rulka, M. P. Scott, and J. A. Kennison, 1991, The arflike gene encodes an essential GTP-binding protein in *Drosophila*: *Proc Natl Acad Sci U S A*, v. 88, p. 3120-4.
- Tanos, B. E., H. J. Yang, R. Soni, W. J. Wang, F. P. Macaluso, J. M. Asara, and M. F. Tsou, 2013, Centriole distal appendages promote membrane docking, leading to cilia initiation: *Genes Dev*, v. 27, p. 163-8.
- Taschner, M., and E. Lorentzen, 2016, The Intraflagellar Transport Machinery: *Cold Spring Harb Perspect Biol*, v. 8.
- Taylor, S. S., and E. Radzio-Andzelm, 1994, Three protein kinase structures define a common motif: *Structure*, v. 2, p. 345-55.
- Tetley, L., and K. Vickerman, 1985, Differentiation in *Trypanosoma brucei*: host-parasite cell junctions and their persistence during acquisition of the variable antigen coat: *J Cell Sci*, v. 74, p. 1-19.
- Tian, G., A. Bhamidipati, N. J. Cowan, and S. A. Lewis, 1999, Tubulin folding cofactors as GTPase-activating proteins. GTP hydrolysis and the assembly of the alpha/beta-tubulin heterodimer: *J Biol Chem*, v. 274, p. 24054-8.
- Tian, G., Y. Huang, H. Rommelaere, J. Vandekerckhove, C. Ampe, and N. J. Cowan, 1996, Pathway leading to correctly folded beta-tubulin: *Cell*, v. 86, p. 287-96.
- Tian, G., S. A. Lewis, B. Feierbach, T. Stearns, H. Rommelaere, C. Ampe, and N. J. Cowan, 1997, Tubulin subunits exist in an activated conformational state generated and maintained by protein cofactors: *J Cell Biol*, v. 138, p. 821-32.
- Tibbetts, R. S., J. L. Jensen, C. L. Olson, F. D. Wang, and D. M. Engman, 1998, The DnaJ family of protein chaperones in *Trypanosoma cruzi*: *Mol Biochem Parasitol*, v. 91, p. 319-26.
- Tobin, J. L., and P. L. Beales, 2009, The nonmotile ciliopathies: *Genet Med*, v. 11, p. 386-402.
- Tolson, D. L., A. Jardim, L. F. Schnur, C. Stebeck, C. Tuckey, R. P. Beecroft, H. S. Teh, R. W. Olafson, and T. W. Pearson, 1994, The kinetoplastid membrane protein 11 of *Leishmania donovani* and African trypanosomes is a potent stimulator of T-lymphocyte proliferation: *Infect Immun*, v. 62, p. 4893-9.
- Tran, P. V., C. J. Haycraft, T. Y. Besschetnova, A. Turbe-Doan, R. W. Stottmann, B. J. Herron, A. L. Chesebro, H. Qiu, P. J. Scherz, J. V. Shah, B. K. Yoder, and D. R. Beier, 2008, THM1 negatively modulates mouse sonic hedgehog signal transduction and affects retrograde intraflagellar transport in cilia: *Nat Genet*, v. 40, p. 403-10.
- Tsai, M. Y., G. Morfini, G. Szebenyi, and S. T. Brady, 2000, Release of kinesin from vesicles by hsc70 and regulation of fast axonal transport: *Mol Biol Cell*, v. 11, p. 2161-73.

- Tschopp, F., F. Charriere, and A. Schneider, 2011, In vivo study in *Trypanosoma brucei* links mitochondrial transfer RNA import to mitochondrial protein import: *EMBO Rep*, v. 12, p. 825-32.
- Týč, J., M. M. Klingbeil, and J. Lukeš, 2015, Mitochondrial heat shock protein machinery hsp70/hsp40 is indispensable for proper mitochondrial DNA maintenance and replication: *MBio*, v. 6.
- UniProtKB-P36405 (ARL3\_HUMAN) <http://www.uniprot.org/uniprot/P36405>.
- Urbaniak, M. D., M. L. Guther, and M. A. Ferguson, 2012, Comparative SILAC proteomic analysis of *Trypanosoma brucei* bloodstream and procyclic lifecycle stages: *PLoS One*, v. 7, p. e36619.
- Urbaniak, M. D., D. M. Martin, and M. A. Ferguson, 2013, Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of *Trypanosoma brucei*: *J Proteome Res*, v. 12, p. 2233-44.
- Urbaniak, M. D., T. Mathieson, M. Bantscheff, D. Eberhard, R. Grimaldi, D. Miranda-Saavedra, P. Wyatt, M. A. Ferguson, J. Frearson, and G. Drewes, 2012, Chemical proteomic analysis reveals the drugability of the kinome of *Trypanosoma brucei*: *ACS Chem Biol*, v. 7, p. 1858-65.
- Uversky, V. N., 2002, What does it mean to be natively unfolded?: *Eur J Biochem*, v. 269, p. 2-12.
- Valente, E. M., C. V. Logan, S. Mougou-Zerelli, J. H. Lee, J. L. Silhavy, F. Brancati, M. Iannicelli, L. Travaglini, S. Romani, B. Illi, M. Adams, K. Szymanska, A. Mazzotta, J. E. Lee, J. C. Tolentino, D. Swistun, C. D. Salpietro, C. Fede, S. Gabriel, C. Russ, K. Cibulskis, C. Sougnez, F. Hildebrandt, E. A. Otto, S. Held, B. H. Diplas, E. E. Davis, M. Mikula, C. M. Strom, B. Ben-Zeev, D. Lev, T. L. Sagie, M. Michelson, Y. Yaron, A. Krause, E. Boltshauser, N. Elkhartoufi, J. Roume, S. Shalev, A. Munnich, S. Saunier, C. Inglehearn, A. Saad, A. Alkindy, S. Thomas, M. Vekemans, B. Dallapiccola, N. Katsanis, C. A. Johnson, T. Attié-Bitach, and J. G. Gleeson, 2010, Mutations in TMEM216 perturb ciliogenesis and cause Joubert, Meckel and related syndromes: *Nat Genet*, v. 42, p. 619-25.
- Vaughan, S., 2010, Assembly of the flagellum and its role in cell morphogenesis in *Trypanosoma brucei*: *Curr Opin Microbiol*, v. 13, p. 453-8.
- Vaughan, S., and K. Gull, 2003, The trypanosome flagellum: *J Cell Sci*, v. 116, p. 757-9.
- Vaughan, S., and K. Gull, 2015, Basal body structure and cell cycle-dependent biogenesis in *Trypanosoma brucei*: *Cilia*, v. 5, p. 5.
- Vaughan, S., L. Kohl, I. Ngai, R. J. Wheeler, and K. Gull, 2008, A repetitive protein essential for the flagellum attachment zone filament structure and function in *Trypanosoma brucei*: *Protist*, v. 159, p. 127-36.
- Veland, I. R., A. Awan, L. B. Pedersen, B. K. Yoder, and S. T. Christensen, 2009, Primary cilia and signaling pathways in mammalian development, health and disease.: *Nephron Physiol*, v. 111, p. p39-53.
- Veltel, S., A. Kravchenko, S. Ismail, and A. Wittinghofer, 2008, Specificity of Arl2/Arl3 signaling is mediated by a ternary Arl3-effector-GAP complex: *FEBS Lett*, v. 582, p. 2501-7.
- Veltel, S., R. Gasper, E. Eisenacher, and A. Wittinghofer, 2008, The retinitis pigmentosa 2 gene product is a GTPase-activating protein for Arf-like 3: *Nat Struct Mol Biol*, v. 15, p. 373-80.
- Vertommen, D., J. Van Roy, J. P. Szikora, M. H. Rider, P. A. Michels, and F. R. Opperdoes, 2008, Differential expression of glycosomal and mitochondrial proteins in the two major life-cycle stages of *Trypanosoma brucei*: *Mol Biochem Parasitol*, v. 158, p. 189-201.
- Vickerman, K., 1969, On the surface coat and flagellar adhesion in trypanosomes.: *J Cell Sci*, v. 5, p. 163-93.

- Vickerman, K., 1985, Developmental cycles and biology of pathogenic trypanosomes: *Br Med Bull*, v. 41, p. 105-14.
- Vidilaseris, K., J. Lesigang, B. Morriswood, and G. Dong, 2015, Assembly mechanism of *Trypanosoma brucei* BILBO1 at the flagellar pocket collar: *Commun Integr Biol*, v. 8, p. e992739.
- Vincensini, L., T. Blisnick, and P. Bastin, 2011, 1001 model organisms to study cilia and flagella.: *Biol Cell*, v. 103, p. 109-30.
- Vincze, O., N. Tökési, J. Oláh, E. Hlavanda, A. Zotter, I. Horváth, A. Lehotzky, L. Tirián, K. F. Medzihradszky, J. Kovács, F. Orosz, and J. Ovádi, 2006, Tubulin polymerization promoting proteins (TPPPs): members of a new family with distinct structures and functions: *Biochemistry*, v. 45, p. 13818-26.
- Vinogradova, T., P. M. Miller, and I. Kaverina, 2009, Microtubule network asymmetry in motile cells: role of Golgi-derived array: *Cell Cycle*, v. 8, p. 2168-74.
- Vos, M. J., J. Hageman, S. Carra, and H. H. Kampinga, 2008, Structural and functional diversities between members of the human HSPB, HSPH, HSPA, and DNAJ chaperone families: *Biochemistry*, v. 47, p. 7001-11.
- Walther, Z., M. Vashishtha, and J. L. Hall, 1994, The *Chlamydomonas* FLA10 gene encodes a novel kinesin-homologous protein: *J Cell Biol*, v. 126, p. 175-88.
- Wang, Y., A. P. McMahon, and B. L. Allen, 2007, Shifting paradigms in Hedgehog signaling: *Curr Opin Cell Biol*, v. 19, p. 159-65.
- Wang, Z., Z. C. Fan, S. M. Williamson, and H. Qin, 2009, Intraflagellar transport (IFT) protein IFT25 is a phosphoprotein component of IFT complex B and physically interacts with IFT27 in *Chlamydomonas*: *PLoS One*, v. 4, p. e5384.
- Waters, A. M., and P. L. Beales, 2011, Ciliopathies: an expanding disease spectrum.: *Pediatr Nephrol*, v. 26, p. 1039-56.
- Wei, Q., K. Ling, and J. Hu, 2015, The essential roles of transition fibers in the context of cilia: *Curr Opin Cell Biol*, v. 35, p. 98-105.
- Wei, Q., Q. Xu, Y. Zhang, Y. Li, Q. Zhang, Z. Hu, P. C. Harris, V. E. Torres, K. Ling, and J. Hu, 2013, Transition fibre protein FBF1 is required for the ciliary entry of assembled intraflagellar transport complexes: *Nat Commun*, v. 4, p. 2750.
- Wei, Q., Y. Zhang, Y. Li, Q. Zhang, K. Ling, and J. Hu, 2012, The BBSome controls IFT assembly and turnaround in cilia: *Nat Cell Biol*, v. 14, p. 950-7.
- Westlake, C. J., L. M. Baye, M. V. Nachury, K. J. Wright, K. E. Ervin, L. Phu, C. Chalouni, J. S. Beck, D. S. Kirkpatrick, D. C. Slusarski, V. C. Sheffield, R. H. Scheller, and P. K. Jackson, 2011, Primary cilia membrane assembly is initiated by Rab11 and transport protein particle II (TRAPP II) complex-dependent trafficking of Rabin8 to the centrosome: *Proc Natl Acad Sci U S A*, v. 108, p. 2759-64.
- Wheeler, R. J., N. Scheumann, B. Wickstead, K. Gull, and S. Vaughan, 2013, Cytokinesis in *Trypanosoma brucei* differs between bloodstream and tsetse trypomastigote forms: implications for microtubule-based morphogenesis and mutant analysis: *Mol Microbiol*, v. 90, p. 1339-55.
- Wicks, S. R., C. J. de Vries, H. G. van Luenen, and R. H. Plasterk, 2000, CHE-3, a cytosolic dynein heavy chain, is required for sensory cilia structure and function in *Caenorhabditis elegans*: *Dev Biol*, v. 221, p. 295-307.
- Wickstead, B., K. Ersfeld, and K. Gull, 2002, Targeting of a tetracycline-inducible expression system to the transcriptionally silent minichromosomes of *Trypanosoma brucei*: *Mol Biochem Parasitol*, v. 125, p. 211-6.
- Wiese, M., Morris, A., & Grant, K. M., 2009, [Trypanosomatid protein kinases as potential drug targets](#). In P. M. Selzer (Ed.), *Antiparasitic and antibacterial drug discovery : from molecular targets to drug candidates*.:pp. 227-247, Drug Discovery in Infectious Diseases.

- Williams, C. L., C. Li, K. Kida, P. N. Inglis, S. Mohan, L. Semenec, N. J. Bialas, R. M. Stupay, N. Chen, O. E. Blacque, B. K. Yoder, and M. R. Leroux, 2011, MKS and NPHP modules cooperate to establish basal body/transition zone membrane associations and ciliary gate function during ciliogenesis: *J Cell Biol*, v. 192, p. 1023-41.
- Williams, C. L., S. V. Masyukova, and B. K. Yoder, 2010, Normal ciliogenesis requires synergy between the cystic kidney disease genes MKS-3 and NPHP-4: *J Am Soc Nephrol*, v. 21, p. 782-93.
- Williams, C. L., J. C. McIntyre, S. R. Norris, P. M. Jenkins, L. Zhang, Q. Pei, K. Verhey, and J. R. Martens, 2014, Direct evidence for BBSome-associated intraflagellar transport reveals distinct properties of native mammalian cilia: *Nat Commun*, v. 5, p. 5813.
- Williams, C. L., M. E. Winkelbauer, J. C. Schafer, E. J. Michaud, and B. K. Yoder, 2008, Functional redundancy of the B9 proteins and nephrocystins in *Caenorhabditis elegans* ciliogenesis: *Mol Biol Cell*, v. 19, p. 2154-68.
- Williams, S., L. Saha, U. K. Singha, and M. Chaudhuri, 2008, *Trypanosoma brucei*: differential requirement of membrane potential for import of proteins into mitochondria in two developmental stages: *Exp Parasitol*, v. 118, p. 420-33.
- Wolburg, H., S. Mogk, S. Acker, C. Frey, M. Meinert, C. Schönfeld, M. Lazarus, Y. Urade, B. K. Kubata, and M. Duszenko, 2012, Late stage infection in sleeping sickness: *PLoS One*, v. 7, p. e34304.
- Won, J., C. Marín de Esvikova, R. S. Smith, W. L. Hicks, M. M. Edwards, C. Longo-Guess, T. Li, J. K. Naggert, and P. M. Nishina, 2011, NPHP4 is necessary for normal photoreceptor ribbon synapse maintenance and outer segment formation, and for sperm development: *Hum Mol Genet*, v. 20, p. 482-96.
- Woodward, R., and K. Gull, 1990, Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*: *J Cell Sci*, v. 95 ( Pt 1), p. 49-57.
- Wright, K. J., L. M. Baye, A. Olivier-Mason, S. Mukhopadhyay, L. Sang, M. Kwong, W. Wang, P. R. Pretorius, V. C. Sheffield, P. Sengupta, D. C. Slusarski, and P. K. Jackson, 2011, An ARL3-UNC119-RP2 GTPase cycle targets myristoylated NPHP3 to the primary cilium: *Genes Dev*, v. 25, p. 2347-60.
- Wu, J., T. Liu, Z. Rios, Q. Mei, X. Lin, and S. Cao, 2016, Heat Shock Proteins and Cancer: *Trends Pharmacol Sci*
- Yaffe, M. B., G. W. Farr, D. Miklos, A. L. Horwich, M. L. Sternlicht, and H. Sternlicht, 1992, TCP1 complex is a molecular chaperone in tubulin biogenesis: *Nature*, v. 358, p. 245-8.
- Yakulov, T. A., T. Yasunaga, H. Ramachandran, C. Engel, B. Müller, S. Hoff, J. Dengjel, S. S. Lienkamp, and G. Walz, 2015, Anks3 interacts with nephronophthisis proteins and is required for normal renal development: *Kidney Int*, v. 87, p. 1191-200.
- Yam, A. Y., Y. Xia, H. T. Lin, A. Burlingame, M. Gerstein, and J. Frydman, 2008, Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies: *Nat Struct Mol Biol*, v. 15, p. 1255-62.
- Yamakami, M., T. Yoshimori, and H. Yokosawa, 2003, Tom1, a VHS domain-containing protein, interacts with tollip, ubiquitin, and clathrin: *J Biol Chem*, v. 278, p. 52865-72.
- Yamamoto, R., M. Hirono, and R. Kamiya, 2010, Discrete PIH proteins function in the cytoplasmic preassembly of different subsets of axonemal dyneins: *J Cell Biol*, v. 190, p. 65-71.
- Yang, C., H. A. Owen, and P. Yang, 2008, Dimeric heat shock protein 40 binds radial spokes for generating coupled power strokes and recovery strokes of 9 + 2 flagella: *J Cell Biol*, v. 180, p. 403-15.
- Yang, C., M. M. Compton, and P. Yang, 2005, Dimeric novel HSP40 is incorporated into the radial spoke complex during the assembly process in flagella: *Mol Biol Cell*, v. 16, p. 637-48.

- Yang, P., C. Yang, M. Wirschell, and S. Davis, 2009, Novel LC8 mutations have disparate effects on the assembly and stability of flagellar complexes: *J Biol Chem*, v. 284, p. 31412-21.
- Young, P., Q. Deveraux, R. E. Beal, C. M. Pickart, and M. Rechsteiner, 1998, Characterization of two polyubiquitin binding sites in the 26 S protease subunit 5a: *J Biol Chem*, v. 273, p. 5461-7.
- Yang, Z. R., R. Thomson, P. McNeil, and R. M. Esnouf, 2005, RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins: *Bioinformatics*, v. 21, p. 3369-76.
- Ye, X., H. Zeng, G. Ning, J. F. Reiter, and A. Liu, 2014, C2cd3 is critical for centriolar distal appendage assembly and ciliary vesicle docking in mammals: *Proc Natl Acad Sci U S A*, v. 111, p. 2164-9.
- Yee, L. E., F. R. Garcia-Gonzalo, R. V. Bowie, C. Li, J. K. Kennedy, K. Ashrafi, O. E. Blacque, M. R. Leroux, and J. F. Reiter, 2015, Conserved Genetic Interactions between Ciliopathy Complexes Cooperatively Support Ciliogenesis and Ciliary Signaling: *PLoS Genet*, v. 11, p. e1005627.
- Yoon, J. H., J. Qiu, S. Cai, Y. Chen, M. E. Cheetham, B. Shen, and G. P. Pfeifer, 2006, The retinitis pigmentosa-mutated RP2 protein exhibits exonuclease activity and translocates to the nucleus in response to DNA damage: *Exp Cell Res*, v. 312, p. 1323-34.
- Zdobnov, E. M., and R. Apweiler, 2001, InterProScan--an integration platform for the signature-recognition methods in InterPro: *Bioinformatics*, v. 17, p. 847-8.
- Zhang, H., S. Li, T. Doan, F. Rieke, P. B. Detwiler, J. M. Frederick, and W. Baehr, 2007, Deletion of PrBP/delta impedes transport of GRK1 and PDE6 catalytic subunits to photoreceptor outer segments: *Proc Natl Acad Sci U S A*, v. 104, p. 8857-62.
- Zhang, P., J. I. Leu, M. E. Murphy, D. L. George, and R. Marmorstein, 2014, Crystal structure of the stress-inducible human heat shock protein 70 substrate-binding domain in complex with peptide substrate: *PLoS One*, v. 9, p. e103518.
- Zhang, Q., J. Hu, and K. Ling, 2013, Molecular views of Arf-like small GTPases in cilia and ciliopathies: *Exp Cell Res*, v. 319, p. 2316-22.
- Zhang, Y., S. Seo, S. Bhattarai, K. Bugge, C. C. Searby, Q. Zhang, A. V. Drack, E. M. Stone, and V. C. Sheffield, 2014, BBS mutations modify phenotypic expression of CEP290-related ciliopathies: *Hum Mol Genet*, v. 23, p. 40-51.
- Zhao, R., M. Davey, Y. C. Hsu, P. Kaplanek, A. Tong, A. B. Parsons, N. Krogan, G. Cagney, D. Mai, J. Greenblatt, C. Boone, A. Emili, and W. A. Houry, 2005, Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the hsp90 chaperone: *Cell*, v. 120, p. 715-27.
- Zhou, C., L. Cunningham, A. I. Marcus, Y. Li, and R. A. Kahn, 2006, Arl2 and Arl3 regulate different microtubule-dependent processes: *Mol Biol Cell*, v. 17, p. 2476-87.
- Zhu, D., D. J. Dix, and E. M. Eddy, 1997, HSP70-2 is required for CDC2 kinase activity in meiosis I of mouse spermatocytes: *Development*, v. 124, p. 3007-14.
- Zhu, G., P. Zhai, X. He, N. Wakeham, K. Rodgers, G. Li, J. Tang, and X. C. Zhang, 2004, Crystal structure of human GGA1 GAT domain complexed with the GAT-binding domain of Rabaptin5: *EMBO J*, v. 23, p. 3909-17.
- Zotter, A., A. Bodor, J. Oláh, E. Hlavanda, F. Orosz, A. Perczel, and J. Ovádi, 2011, Disordered TPPP/p25 binds GTP and displays Mg<sup>2+</sup>-dependent GTPase activity: *FEBS Lett*, v. 585, p. 803-8.
- Zhou, Q., H. Hu, C. Y. He, and Z. Li, 2015, Assembly and maintenance of the flagellum attachment zone filament in *Trypanosoma brucei*: *J Cell Sci*, v. 128, p. 2361-72.
- Zhou, Q., H. Hu, and Z. Li, 2014, New insights into the molecular mechanisms of mitosis and cytokinesis in trypanosomes: *Int Rev Cell Mol Biol*, v. 308, p. 127-66.

- Zhou, Q., B. Liu, Y. Sun, and C. Y. He, 2011, A coiled-coil- and C2-domain-containing protein is required for FAZ assembly and cell morphology in *Trypanosoma brucei*: *J Cell Sci*, v. 124, p. 3848-58.
- Čajánek, L., and E. A. Nigg, 2014, Cep164 triggers ciliogenesis by recruiting Tau tubulin kinase 2 to the mother centriole: *Proc Natl Acad Sci U S A*, v. 111, p. E2841-50.